الإحصاء

و

الاحتمال

تأليف

الدكتور أنيس إسماعيل كنجو

أستاذ بقسم الإحصاء-كلية العلوم

جامعة الملك سعود

مكتبة العبيكان

ⓗ مكتبة العبيكان، ١٤٢٠هـ

فهرسة مكتبة الملك فهد الوطنية أثناء النشر

كنجو، أنيس

الإحصاء والاحتمال – الرياض.

٤٥٦ص، ١٧X٢٤سم

ردمك: ٤-٦٣٩-٢٠-٩٩٦٠

١- الاحتمالات (رياضيات)	٢- الإحصاء الرياضي

أ- العنوان

ديوي ٥١٩	٢٠/٣٠٨٧

ردمك: ٤-٦٣٩-٢٠-٩٩٦٠	رقم الإيداع: ٢٠/٣٠٨٧

الطبعة الأولى الخاصة بمكتبة العبيكان

١٤٢١هـ / ٢٠٠٠م

طبعة مزيدة ومنقحة

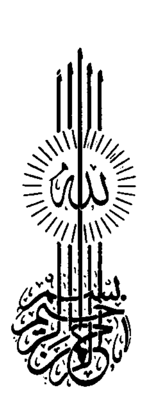
الناشر

مكتبة العبيكان

الرياض ـ العليا ـ تقاطع طريق الملك فهد مع العروبة.

ص.ب: ٦٢٨٠٧ الرياض ١١٥٩٥

هاتف: ٤٦٥٤٤٢٤، فاكس: ٤٦٥٠١٢٩



مقدمة الطبعة الأولى

الحمد لله وحده والصلاة والسلام على نبينا محمد وبعد، فقد شعرت، نتيجة تدريسي لمقرر إحص ١٢١ وغيره من مقررات المستوى الأول في الإحصاء سنة بعد أخرى، بالحاجة الملحّة إلى تأليف كتاب دراسي يغطي بصورة رئيسة منهاج هذا المقرر، ويتضمن عرضا لمبادىء الإحصاء والاحتمال، لا يقتصر على خطوات الطريقة الإحصائية وسُبل حسابها، وإنما يتطرق إلى كنه المسألة الإحصائية وصلتها الحميمة بالاحتمال، فيوضحها بطريقة ميسرة وسهلة خالية قدر الإمكان من اللّبس والغموض، وملتزمة قدر المستطاع بالأمانة العلمية الضرورية، وبالدقة التي يسمح بها مستوى طالب جامعي في سنته الأولى.

والكتاب إذ يعفي المدرس من ضرورة الكتابة المسهبة على السبورة لأفكار المحاضرة، إنما يفسح المجال رحبا لنقاش مستفيض يجتذب انتباه الطلبة وعقولهم، وسير أمثل للمحاضرة، يشارك فيه الطالب مشاركة فعلية في التحليل والاستنتاج، ويُسهم بكل ادراكه وقدرته على التركيز والانتباه في استنباط المفاهيم والتعليق عليها وإبداء ما يدور في ذهنه من تساؤلات حولها؛ وذلك بدلاً من أن يكون آلة تسجيل تنسخ ما يُكتب على السبورة، وربما دون أن يفكر فيما يكتب. ويجد المحاضر نفسه في صراع حقيقي بين رغبته في تغطية المنهج الواسع، بكل ما يحتويه من مفاهيم غنية وجديدة تُطرح على الطالب للمرة الأولى في حياته، وبين رغبته في إعطاء تلك المفاهيم حقها من الشرح والإيضاح، والوصول إلى قناعة الطالب فيها من خلال القياس

والمقارنة، وتحقيق أوسع مشاركة ممكنة للطالب في سير المحاضرة وإبقائه مُستنفراً يقظا بدلا من تركه فريسه سهلة لغفوة النسخ الرتيب.

وفي اعتقادي أن محاضرة الإحصاء بخاصّة تحتاج، إلى نوع من شد الذهن وترويضه. وإبقائه في حالة تحفز، إذ تزخر عادة بمعالجة طيف متعدد الألوان من مشكلات الحياة على اتساعها، وبطريقة تتميز بالخروج على النمطية واللجوء إلى مفاهيم وطرق من التفكير والتطبيق لم يألفها الطالب من قبل. فمع الأسف الشديد، لا تقدم له مراحل ما قبل الجامعة، أي قدر من التدريب في مجال العشوائية أو أي نصيب من الإلفة بالطبيعة التكرارية للمسألة الاحتمالية والإحصائية.

ولما كان يمكن لطلاب هذا المقرر أن يكونوا طلاب رياضيات يستهلون به إعدادهم المتواضع في مجال الإحصاء والاحتمال. أو طلابا من تخصص علوم الأحياء يشكل المقرر بالنسبة لهم ليس بداية المطاف فقط وإنما، في الغالب، خاتمته أيضا. ويكون هؤلاء عادة ممن تجاوزوا في الغالب مستوى السنة الأولى، وقد يكونون من المستويين الثالث والرابع، مما يتيح الفرصة لتزويدهم ببعض الأفكار الأساسية في الاستقراء الإحصائي. فقد حَرِصْت على أن يتضمن الكتاب فقرات منجمة لا تعتبر من صلب المنهج، ولكنها تترك للمدرس إمكانية تزويد طلابه بما يراه مناسباً لهم من هذه الفقرات ومن التمارين الموافقة لها، أو يعتبرها مادة للقراءة والاطلاع فقط. ومع هذه الفقرات يتسع، إلى حد ما، مدى الفائدة من الكتاب، كما تتسع ساحة المستفيدين منه. وينسجم هذا التدبير مع حرصي على أن يشكل الكتاب مرجعا مفيدا لقراء من خارج الاختصاص يتمتعون بدرجة جيدة من النضج الذهني، ولكنهم يفتقرون في تدريبهم السابق لأي معرفة بأوليات الإحصاء.

ولقد توخيت من طريقة العرض خروج الدارس الذي سيكتفي بمقرر واحد في الإحصاء، بفكرة واضحة قدر الإمكان عن طبيعة المسألة الإحصائية ودور الإحتمال فيها. فركزت قدر إمكاني على توضيح ظاهرة الانتظام الإحصائي، ومفهوم العشوائية

والمتغير العشوائي ، والتوزيع الاحتمالي وتفسيره العملي ، والعينة العشوائية ودورها . واستخدمت لغة العينة والمجتمع حيثما أمكن ذلك ، ولم أترك فرصة متاحة للخوض في أوليات الاستقراء الإحصائي إلا اهتبلتها ، مستهدفا الوصول إلى قناعة القارىء عن طريق المناقشة والأمثلة الموجهة والقياس والمقارنة ، معتمدا في ذلك على ما تمليه الفطرة والبداهة وسلامة الإحساس .

ويتضمن الكتاب عددا كبيرا من الأمثلة المحلولة والتمارين . وقد تطلب جمعها وترتيبها جهدا إضافيا خاصا ، فهي ليست تكرارا مملا للفكرة نفسها ، وعلى الوتيرة والمستوى نفسيهما ، وإنما تطرق أفكارا متنوعة مستوحاة من واقع الحياة . وتتدرج في مستواها من السهل إلى الصعب . وبعضها يشكل تحديا بسيطا يرحب به الطالب الممتاز.

وإذ أقدم هذا الجهد المتواضع للقارىء العربي أرجو من الله سبحانه وتعالى أن يتقبله مني عملا صالحا فهو من وراء القصد وهو الهادي إلى سواء السبيل .

المؤلف

مقدمة الطبعة الثانية

لاقت الطبعة الأولى بحمد الله وعونه قبولاً ملحوظاً من القراء الكرام. وقد تلقيت من العديد من الزملاء داخل جامعة الملك سعود وخارجها ومن زملاء خارج المملكة ممن اطلعوا على الكتاب أو استخدموه في تدريسهم ما يفيد بأن الكتاب ناجح في عرضه الواضح والمبسط من جهة، والدقيق، في حدود ما يسمح به مستوى الكتاب، من جهة أخرى. ولم أجد أي ضرورة لإدخال تعديلات أو إضافات على محتوى الكتاب في طبعته الثانية باستثناء تصويبات تناولت الأخطاء الطباعية.

أسأل الله أن يكون في هذه الطبعة كل الفائدة التي أتوخاها للقارئ الكريم وأن يتقبل مني هذا الجهد عملاً صالحاً لوجهه الكريم.

المؤلف

أ.د. أنيس كنجو

المحتويات

صفحة

صفحة

صفحة

صفحة

صفحة

صفحة

صفحة

مقدمة الكتاب

لا شك في أن لـدى كل قـارىء إدراكا معينـا لكلمـة «الإحصاء». وتنشأ هـذه المدارك مما اعتـاد عليه عامة النـاس في أيامنا هـذه من إطلاع شبه يومي على معلـومات إحصائية تقدمها النشرات الإحصائية الرسمية للحكومات والهيئات الدولية ومن خلال الصحافة ووسـائل الإعلام المسموعة والمرئيـة. وهي تشير، في الغالب، إلى أن الإحصاء هو نوع من التجميع لقدر كبير من المعلومات الكمية أو الكيفية واختزالها وتقديمها على شكل جـداول أو رسوم أو أشكـال وخطوط بيـانية معبرة وسهلـة التناول والادراك. كما يمكن أن تتضمن حساب مجاميع أو معدلات أو نسب مئوية أو ماشابهها.

وربما كانت إحـدى الفوائد المتـوخاة لمقـرر ابتدائي في الإحصاء هـي التزود بفهم أكثر شمـولا وعمقا ودقة لكلمة «الإحصـاء» بمعناها العلمي المعاصر، فـالفهم السائد لكلمة الإحصاء يندرج، في الواقع، تحت عنـوان «الإحصاء الوصفي». ولكن الإحصاء يلعب اليـوم دورا مزدوجـا إذ يقـدم إلى جـانب الإحصاء الـوصفي طـرقـا للاستقـراء، فنستخلص من البيان الإحصائي نتائج معينة بطريقة تتسم بالموضوعية. ولا شك في أن جانب الاستقـراء الإحصائي هو الجانب الأكثـر إثارة ومدعاة لـلاهتمام، ويشكل اليوم إحدى أهم الأدوات المعـاصرة لاتخاذ قرار أو القيام بتنبؤ في ظروف تخضـع للمصادفة، أي ظروف لا يمكن معها التنبؤ بالنتـائج أو محاولة التعرف على القرار السليم من خلال قوانين علمية معـروفة. ويهدف كتاب ابتدائي كهـذا، فيما يهدف، إلى نقل القارىء إلى مشارف الاستقراء الإحصائي. وإذا كان الكتاب بأكمله لا يطمح في هذا الخصوص إلى أكثر من ذلك، فمن المستحيل على مقدمة مختصرة وسريعة أن تدعي المقدرة على تقديم

فكرة واضحة ودقيقة عن ما هية الاستقراء الإحصائي. ومع ذلك لا بد لنا من إلقاء بعض الضوء على مصطلحين أساسيين في علم الإحصاء، هما المجتمع والعينة. وسنحاول تلمس العناصر الأساسية للمسألة الإحصائية مهتدين في ذلك بنقاط رئيسة تضمنتها نشرة علمية بعنوان: "Careers in Statistics" أصدرتها عام ١٩٦٢م أكبر هيئتين علميتين إحصائيتين في الولايات المتحدة هما:

"The American Statistical Association"

"The Institute of Mathematical Statistics"

يهدف الإحصاء باعتباره فرعا من فروع الطريقة العلمية إلى دراسة خصائص عددية للمجتمعات. ولكن ماذا نقصد بمصطلح «مجتمع»؟

في معظم الأبحاث العلمية لا ينصب الاهتمام على البيان الإحصائي المدروس وإنما يكتسب البيان أهميته من كونه ممثلا لمجموعة أكبر من المعلومات الإحصائية يشكل البيان المدروس جزءا منها.

وعلى سبيل المثال إذا سألنا مائة طالب من طلاب كلية العلوم عن رأيهم في الدورة المكثفة في اللغة الإنجليزية، فإن آراء الطلاب المائة لذاتها ليس لها أهمية كبيرة، وإنما تأتي أهميتها من كونها مؤشرا للرأي السائد بين مجموعة أكبر بكثير من الطلبة هم جميع طلبة كلية العلوم. ونصطلح في الإحصاء على تسمية الطلاب المائة «عينة» ومجموعة طلبة كلية العلوم «المجتمع». والدراسة تهدف أول ما تهدف إلى التعرف على الرأي السائد بين طلبة كلية العلوم إزاء الدورة المكثفة. أي أن هدف الدراسة هو المجتمع.

وإليك مثال ثان. لنفرض أن عدد المستجدين في الجامعة هو خمسة آلاف طالب، وأن باحثا يرغب في معرفة مجموع أوزان المستجدين. فالمجتمع هنا هو كافة المستجدين في الجامعة ويمكن للباحث أن يقوم بوزنهم واحدا فآخر ويصل إلى ما

يريد، كما يمكنه اتباع طريقة أخرى، فيختار مائتي طالب، مثلا، ويقيس أوزانهم، ومن هذه القياسات يحاول تقدير الوزن الكلي لجميع الطلبة المستجدين. ويشكل الطلاب المائتان الذين إختارهم عينة من مجتمع المستجدين، ويسمى وزن الطالب، قياسا أو ملاحظة أو مشاهدة.

ومثال ثالث. لنفترض أن باحثا في العلوم الطبية يرغب في تثمين دواء جديد لمرض معين. وقد طبق المعالجة الجديدة على عشرين مريضا، فمن وجهة نظر الباحث لا يشكل المرضى العشرون المجتمع الذي يهدف إلى دراسته، وإنما يشكلون عينة منه فقط. وهو لا يهتم بنتائج المعالجة بين هؤلاء المرضى العشرين لذاتهم وإنما يهمه معرفة مدى نجاح المعالجة من أجل أي مصاب بذلك المرض. والمجتمع الذي يهمه هو إذا مجتمع جميع المصابين بهذا المرض ويمكنهم تلقي العلاج، سواء من كان منهم موجودا الآن ومن سيوجد في المستقبل. والمجتمع هنا هو نوع من المجتمع التصوري، إذ لا وجود له في الواقع المحسوس، ومع ذلك فهو المجتمع الذي ينصب عليه الاهتمام، لأن الباحث يريد تثمين معالجته وهي تطبق على المصابين بهذا المرض بصورة عامة، وليس على المرضى العشرين الذي يشكلون العينة.

وعندما يكرر باحث في العلوم الفيزيائية، مثلا، تجربة قياس ثابت فيزيائي معين، عشر مرات، فنصطلح على اعتبار القياسات العشرة، التي يحصل عليها، عيّنة من مجتمع تصوري يتضمن جميع القياسات التي كان سيحصل عليها الباحث لو أنه استمر في تكرار تجربته عددا لا نهائيا من المرات. والمجتمع في هذه الحالة تصوري وغير محدود (لا نهائي).

وبصورة عامة، يمكن القول إن المجتمع هو جملة الأشياء أو العناصر التي تشكل هدف الدراسة، أما العينة فهي الجزء من المجتمع الذي يخضع بالفعل للدراسة.

والسؤال الذي يفرض نفسه الآن هو: لماذا لا تتناول الدراسة المجتمع كله؟ وإذا كنا نريد معلومات تتعلق بالمجتمع كله فلماذا نكتفي بجمع معلومات من عينة منه

فقط؟ وللإجابة نقول إن المجتمعات غالبا ما تكون من الضخامة بحيث يكون إخضاع كل عنصر فيها للدراسة نوعا من المستحيل. وحتى عندما يكون ذلك ممكنا من الناحية النظرية، على الأقل، فإن ما تتطلبه الدراسة من جهود وزمن ونفقات طائلة تجعلها من الناحية الواقعية أمرا غير عملي البتة. لا بل قد تقدم لنا دراسة متأنية ودقيقة للعينة، من المعلومات، أفضل مما تقدمه دراسة تتناول المجتمع كله، ولكنها دراسة تنقصها الدقة وتسودها الفوضى. ففي مثال المستجدين يمكن للباحث أن يقوم بوزن كل طالب من طلاب العينة المائتين بدقة، ولكنه إذا حاول الحصول على أوزان المستجدين بآلافهم الخمسة فقد يضطر إلى الاقتناع بتوجيه سؤال إلى الطالب عن وزنه ويكتفي بتسجيل الإجابة، وقد يكون الجواب بعيدا كل البعد عن الدقة. أما عندما يكون المجتمع تصوريا فنجد أنفسنا ملزمين بالإعتماد على عينة، ولا خيار لنا في ذلك.

وفي الإحصاء نعتمد عادة على عينات نختارها عشوائيا ونسميها عينات عشوائية. ولكن ماذا نقصد بكلمة عشوائية؟ ولماذا نريد للعينة أن تكون عشوائية؟ لنفرض، على سبيل المثال، أننا نريد تقديم جائزة لطالب نختاره عشوائيا من فصل يتضمن ثلاثين طالبا، فكيف يتم مثل هذا الاختيار العشوائي؟ إن أي طريقة اختيار نقتنع جميعا أنها خالية تماما من التحيز لمصلحة طالب دون آخر هي طريقة يمكن أن توصف بالعشوائية. لنقم، مثلا، بتسجيل اسم كل طالب على قطعة واحدة من الورق، ثم نطوي هذه الأوراق ونضعها في قبعة، ثم لنخلطها جيدا قبل أن نختار واحدة منها، دون النظر إلى القبعة، ونقدم الجائزة للطالب الذي كُتب اسمه عليها. وسنوافق على وصف هذه الطريقة بأنها عشوائية إذا لم تتضمن أي عمل أو تصرف يمكن أن يساعد على التحيز في الاختيار لمصلحة طالب أو طلاب معينين. فالقطع من الورق يجب أن تكون من الحجم والملمس ونوع الورق نفسه، وتطوى بالطريقة نفسها بحيث تكون متماثلة في كل شيء باستثناء الإسم الذي كتب عليها، وبحيث يمتنع على من يختار الاستفادة بأي صورة من الصور من حاسة اللمس أو النظر. ولا بد أن تخلط الأوراق خلطا جيدا قبل الشروع في اختيار إحداها. وبالمعنى الاصطلاحي للكلمة تطلق كلمة «عشوائي» على أي طريقة اختيار لا هدف لها ولا غاية. ونتحدث عادة عن اختيار أسماء من قبعة عشوائيا، وعن اختيار سنابل قمح عشوائيا من حقل قمح، واختيار أسرة عشوائيا من مجتمع من الأسر في مدينة، الخ. ونعني بذلك أن يتم

الاختيار بفعل المصادفة البحتة وأن تتاح الفرصة نفسها عند كل سحب لكل عنصر من عناصر المجتمع الذي نسحب منه.

وسبب اعتمادنا على العشوائية في علم الإحصاء هو أنها تسمح بتطبيق الطرق الرياضية بسهولة، مما يؤدي إلى استخلاص نتائج تتعلق بالمجتمع بطريقة تتسم بالموضوعية. والجدير بالذكر أنها تقي من آثار التحيز الشخصي، إذ لا يجوز بالطبع أن نترك للباحث الحرية في اختيار عينته، فقد يختارها عندئذ بصورة متحيزة تدعم نظريته.

وفي المجتمعات المحدودة التي يمكن ترقيم عناصرها من 1 إلى عدد محدود N، حيث N عدد الوحدات أو العناصر في المجتمع، توجد جداول للأرقام العشوائية هي جداول كل رقم فيها اختير عشوائيا من بين الأرقام 9 ,... ,2 ,1 ,0. وقد أعدت بحيث يكون لكل رقم من هذه الأرقام الفرصة نفسها في أن يكون الرقم المسحوب وذلك عند كل سحب. ومن بين الجداول الأكثر انتشارا نجد تلك التي نشرتها مؤسسة راند (RAND) عام ١٩٥٥م، وتتضمن مليون رقم. ويعرض الجدول (ا) التالي ألف رقم عشوائي للتوضيح.

وعند استخدام هذه الجداول لاختيار عينة عشوائية بسيطة تكون الخطوة الأولى هي ترقيم الوحدات في المجتمع من 1 إلى N، حيث N عدد وحدات المعاينة في المجتمع، وإذا كان الرقم الأول (من اليسار) للعدد N بين 5 و 9 تكون الطريقة التالية مناسبة. فلنفرض للتوضيح أن المجتمع يتضمن 528 وحدة، أي N = 528، ونريد عينة عشوائية بسيطة حجمها 10 وحدات، فنختار، لا على التعيين، أحد أعمدة الجدول الخمسين، وليكن مثلا العمود 25 ونأخذ العمودين التاليين له وهما العمود 26 والعمود 27، فتعطينا الأرقام المتجاورة (الواقعة على السطر نفسه) من الأعمدة الثلاثة عددا من ثلاثة أرقام. نستعرض هذه الأعداد من الأعلى إلى الأسفل، ونختار الأعداد المتميزة العشرة الأولى الواقعة بين 001 و 528 فنجد 36، 509، 364، 417، 348، 127، 149، 186، 290، و162. وتكون العينة العشوائية المطلوبة هي الوحدات التي تحمل هذه الأرقام. (من أجل العددين الأخيرين قفزنا إلى الأعمدة 30، 31، 32). وعند اختيار عينات مختلفة يستحسن تغيير النقطة التي نبدأ عندها في الجدول من عينة إلى أخرى.[1]

---

(١) انظر كتاب "Sampling Techniques" لمؤلفه W. Cochran، الطبعة الثالثة، صفحة 19.

جدول (١) يوضح ألف رقم عشوائي

	00-04	05-09	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
00	54463	22662	65905	70639	79365	67382	29085	69831	47058	08186
01	15389	85205	18850	39226	42249	90669	96325	23248	60933	26927
02	85941	40756	82414	02015	13858	78030	16269	65978	01385	15345
03	61149	69440	11286	88218	58925	03638	52862	62733	33451	77455
04	05219	81619	10651	67079	92511	59888	84502	72095	83463	75577
05	41417	98326	87719	92294	46614	50948	64886	20002	97365	30976
06	28357	94070	20652	35774	16249	75019	21145	05217	47286	76305
07	17783	00015	10806	83091	91530	36466	39981	62481	49177	75779
08	40950	84820	29881	85966	62800	70326	84740	62660	77379	90279
09	82995	64157	66164	41180	10089	41757	78258	96488	88629	37231
10	96754	17676	55659	44105	47361	34833	86679	23930	53249	27083
11	34357	88040	53364	71726	45690	66334	60332	22554	90600	71113
12	06318	37403	49927	57715	50423	67372	63116	48888	21505	80182
13	62111	52820	07243	79931	89292	84767	85693	73947	22278	11551
14	47534	09243	67879	00544	23410	12740	02540	54440	32949	13491
15	98614	75993	84460	62846	59844	14922	48730	73443	48167	34770
16	24856	03648	44898	09351	98795	18644	39965	71058	90368	44104
17	96887	12479	80621	66223	86085	78285	02432	53342	42846	94771
18	90801	21472	42815	77408	37390	76766	52615	32141	30286	18106
19	55165	77312	83666	36028	28420	70219	81369	41943	47366	41067

وطريقة اختيار عينة عشوائية تصبح مسألة معقدة عندما لا تتوفر قائمة بوحدات المجتمع، أو لا يمكن ترقيم تلك الوحدات. وقليل من التأمل في كيفية اختيار عينة عشوائية من المنازل في مدينة كبيرة، أو من الأشجار في غابة، أو من السمك في بحيرة، أو من المرضى المصابين بعلة معينة، ينبغي أن يقنعنا بأن الصعوبات عديدة ومتنوعة. ويبقى ابتكار طريقة مناسبة تضمن عشوائية العينة أمرا مطلوبا من الباحثين العاملين في حقل المعاينة الإحصائية. وتقنيات اختيار عينة أو ما يسمى بتقنية المعاينة الإحصائية هو عنوان بارز وضخم في أدبيات الإحصاء.

وعندما يكون المجتمع تصوريا وغير محدود، كما في مثال تجربة قياس ثابت فيزيائي، نصطلح على اعتبار التكرارات الـn الأولى للتجربة عينة عشوائية حجمها n من ذلك المجتمع، شريطة أن تتم التكرارات تحت الشروط والظروف نفسها وأن يكون بعضها مستقلا عن بعض.

العناصر الرئيسة لمسألة إحصائية

سنحاول الآن عرض مزيد من الأمثلة نستعرض من خلالها أشكالا من المسائل الإحصائية ونتلمس منها العناصر الرئيسة في مسألة إحصائية. ونسوقها هنا على سبيل المثال لا الحصر، ويحتاج فهم الطرق المتبعة في هذه المسائل إلى العديد من المقررات في نظرية الإحصاء وتطبيقاتها.

١) تقوم إدارة مصنع بتفتيش شحنات البضاعة الخام الواردة إلى المصنع وعلى أساس هذا التفتيش تتخذ قرارا بقبول البضاعة أو رفضها و إعادتها إلى الممول. ويمكن أن يتضمن التفتيش اختيار عينة عشوائية من عشرين وحدة، مثلا، من الشحنة الواردة وفحصها بدقة للوصول إلى عدد الوحدات غير المقبولة من بينها. وعلى أساس هذا العدد يُتخذ قرار برفض الشحنة أو قبولها. إن طريقة اختيار العينة وتحديد حجمها وطريقة اتخاذ القرار هي كلها مسائل إحصائية.

٢) يتوقف إنتاج منشأة للصناعات الكيميائية على عوامل عدة. ويمكن وضع معادلة تنبؤ تربط بين الإنتاج وبين مستويات هذه العوامل وذلك بعد ملاحظة وتسجيل قيمة الإنتاج وقيم هذه العوامل لفترة زمنية معينة. ولكن كيف نضع معادلة تنبؤ جيدة؟ وعند استخدام المعادلة للتنبؤ بالإنتاج لن يكون التنبؤ مساويا للإنتاج الفعلي، بل سيكون هناك دائما فارق أو حيدان بين قيمة التنبؤ والقيمة الفعلية، فكيف يمكن التحكم بهذا الفرق أو الحيدان ووضع حدود دنيا وعليا لمقدار الحيدان؟ وأخيرا ما هي العوامل الأكثر أهمية في عملية الإنتاج؟ وهذه جميعها مسائل إحصائية. ونواجه مثل هذه المسائل في العديد من ميادين المعرفة نذكر منها، على سبيل المثال لا الحصر، العلوم السلوكية (علم التربية، علم الاجتماع، علم النفس. . .)، العلوم الحيوية، العلوم الهندسية والصناعية، العلوم الزراعية، العلوم الاقتصادية، الخ.

٣) يدعي فريق من الباحثين في العلوم الطبية أنهم توصلوا إلى لقاح جديد فعّال في مجال الوقاية من الزكام. فهل ترفض دعواهم أم يجاز تصنيع اللقاح وطرحه للاستهلاك

على نطاق واسع؟ ولنفرض للتبسيط أن اللقاح أعطي لعشرة أشخاص روقبوا طيلة فصل الشتاء وقد جانب الزكام ثمانية منهم فهل يكون اللقاح فعالا؟ إن تصميم التجربة واختيار الأشخاص وتحليل النتائج للوصول إلى قرار حول صلاحية اللقاح هي جميعها مسائل إحصائية. وكم من المواقف المشابهة يتعرض لها الباحثون يوميا ويتعين عليهم الحكم أو اتخاذ قرار بين بديلين مطروحين!

٤) لنفرض أننا قدمنا موضوعا معينا بطريقتين مختلفتين في التدريس إلى مجموعتين من الطلاب لا تتفوق إحداهما على الأخرى في مقدرتها العامة. ثم حصلنا في نهاية الفترة الدراسية على قياس معين لما أنجزته كل من الطريقتين، وعلى أساس من هذه المعلومات نتساءل عما إذا كانت نتائج التجربة تقدم دلالة كافية على تفوق إحدى الطريقتين على الأخرى؟

٥) وعلى مستوى أعم يمكن أن تتطرق الدراسة إلى عدد من المعالجات التي يعتقد أن لها أثرها على ناتج نهائي. فلنفرض ثلاثة أنواع من الأسمدة تختلف في تركيبها من حيث نسبة الأزوت والبوتاس والفوسفات في كل منها. ويمكن تطبيقها في حقول القمح بثلاثة مستويات مختلفة، فنرش مساحة معينة من الأرض، بعدد من الكيلوغرامات أو ضعفي ذلك، أو ثلاثة أضعاف ذلك، كما يمكن توقيت رش السماد في فترتين مختلفتين، فأي الأسمدة، وأي مستويات التطبيق، وأي توقيت للرش أفضل بالنسبة لزيادة إنتاج القمح؟ وإذا كان المستوى الأعلى هو الأجود، مثلا، فهل هناك مجال لمزيد من تحسين الإنتاج من خلال رفع مستوى التطبيق؟ وإلى أي حد يمكن أن نمضي في مثل هذه العملية؟

وتختلف الأمثلة السابقة في طبيعتها ودرجة تعقيدها. إلا أنها تشترك في أن كلا منها ينطوي على تنبؤ أو اتخاذ قرار. بالإضافة إلى أننا في كل من هذه الأمثلة قد أخذنا عينة من كيان أكبر بكثير يدعى المجتمع. والجدير بالذكر أن نتائج العينة لا تُملي علينا القرار أو التنبؤ، فعند مقارنة طريقتين مختلفتين في التدريس، مثلا، لا نلجأ إلى المقارنة الظاهرية المباشرة بين أداء الطريقتين لتفضيل إحداهما على الأخرى، وإنما نلجأ إلى طرق

إحصائية تسمح لنا باتخاذ القرار في سياق العينة التي بين أيدينا وجميع العينات الأخرى الممكنة من الحجم نفسه التي كنا سنحصل عليها لو أننا كررنا تجربة أخذ العينة مرة بعد أخرى. ونعتمد هنا اعتمادا حاسما على نظرية الاحتمالات، ويبدو أساسيا إذاً أن نقوم بتحليل البيان الإحصائي الملحوظ ثم نستقرىء، استنادا إلى التحليل، المجتمع الذي جاءت منه العينة. وثمة عنصر أساسي ثالث لا يبدو بوضوح، فالبيان الإحصائي يحوي قدرا معينا من المعلومات عن الخاصة المدروسة من خصائص المجتمع. وقد تم الحصول على هذه المعلومات نتيجة جهد مبذول كلف مالا ووقتا صرفناهما في تجربة معينة. ولا بد أن قدرا معينا من النفقات والجهود سينتج مقادير مختلفة من المعلومات تبعا لطرق تجريبية مختلفة. ولذلك فمن الواجب تصميم التجربة أو تصميم إجراءات أخذ العينة بحيث نحصل على أكبر قدر من المعلومات المطلوبة لقاء نفقة معينة، ونلخص بقولنا إن المسألة الإحصائية تتضمن:

١ ـ تصميم التجربة أو طريقة أخذ العينة وتجميع البيانات.
٢ ـ تحليل البيان الإحصائي الناتج.
٣ ـ الإستناد إلى هذا التحليل للقيام باستقراء المجتمع الذي جاءت منه العينة.



الفصل الأول

التوزيع الوصفي لجملة من القياسات

(١ ـ ١) اختزال بيان إحصائي وجدول التوزيع التكراري

إن البيانات التي نحصل عليها عند القيام بتنفيذ تجربة أو جمع معلومات إحصائية هي قياسات عددية (كمية) أو وصفية. ومهما أوتينا من الدقة وحسن التتبع فلن يقدم لنا استعراض وتأمل هذه القياسات بطريقة مباشرة وبسيطة، وفي بيانات كبيرة الحجم، إلا القليل جدا عن مدلول هذه القياسات وتفسيرها، وتغيرها بالنسبة لبعضها البعض، ومدى هذا التغيّر. وفي الغالب تبرز فوائد جمة من تصنيف القياسات فيما سنسميه توزيعات تكرارية. وهي تسمح لنا بفهم خصائص وصفية وكمية للبيان الإحصائي، وتفسيره، والحكم عليه بطريقة أكثر موضوعية وأسهل تناولا.

مثال (١ ـ ١)

سألنا عشرة من طلاب الأول ثانوي: «هل ستختار الفرع العلمي أو الفرع الأدبي في العام القادم؟»

وكانت الأجوبة كما يلي:

علمي، علمي، أدبي، علمي، أدبي، أدبي، علمي، علمي، أدبي، علمي

ونلاحظ أن الاختيار «علمي» يظهر ست مرات، أي أن تواتر أو تكرار ظهوره هو 6، بينما يتكرر ظهور الاختيار «أدبي» 4 مرات. ويمكننا ترتيب هذه المعلومات في جدول على الشكل التالي:

جدول (١ ـ ١)

الاختيار	علمي	أدبي
التكرار f	6	4

ونرمز للتكرار بالحرف f.

ويسمى هذا الترتيب للمعلومات التي جمعناها توزيعا تكراريا. فهو يوضح كيف تتوزع الأجوبة العشرة بين الاختيارين المطروحين: علمي، أدبي.

مثال (١ ـ ٢)

يتضمن الجدول (١ ـ ٢) قياس مستوى الهيموغلوبين في الدم لتسعين عاملا ممن يعيشون في مناطق ترتفع ارتفاعا شاهقا عن سطح البحر، والقياسات كما وردت في الجدول تمثل بيانا إحصائيا انتظمت فيه القياسات وفقا لترتيب الحصول عليها أثناء إجراء البحث الإحصائي. فالقياس الأول 18.5 هو مستوى الهيموغلوبين عند أول عامل تناولته التجربة، والقياس الثاني 23.3 هو مستوى الهيموغلوبين عند العامل الثاني الذي تناولته التجربة، وهكذا. ولنفرض أن مما نهتم به في تجربة كهذه، معرفة نسبة العمال الذين يقل مستوى الهيموغلوبين لديهم عن 17. فسيكون الحصول على هذه النسبة من البيان الإحصائي الخام كما ورد في الجدول (١ ـ ٢)، أمرا يستهلك الكثير من الوقت والجهد. وأول ما يخطر بالبال هو تنظيم عرض هذه القياسات بحيث يسهل ترتيبها من الأصغر إلى الأكبر. ولهذا الغرض يمكن إقامة جدول كالجدول (١ ـ ٣)، حيث وضعنا في العمود الأول أعدادا متسلسلة تمثل الرقمين الأولين لقياس (مبتدئين من اليسار) وفي العمود الثاني وضعنا الرقم الثالث (وهو الرقم الأخير) لكل قياس حذاء العدد المناسب، وبحيث تمتد، كما يوضح الجدول، في سطر أفقي، وذلك حسب ترتيب ورودها في البيان. وفي العمود الثالث وضعنا عدد القياسات التي انتظمت أو اصطفت في سطر واحد. وسنطلق على هذه العملية عملية تصفيف البيان الإحصائي الوارد في الجدول (١ ـ ٢).

جدول (١ ـ ٢) قياسات مستوى الهيموغلوبين في الدم لتسعين عاملا

18.5	16.3	23.2	19.4	19.5	20.6	22.0	17.8	16.2
23.3	19.7	21.6	24.2	21.4	20.8	19.7	21.1	23.0
21.7	18.4	22.7	20.9	20.5	16.1	16.9	24.8	12.2
17.4	17.8	19.3	17.3	18.3	17.8	17.1	18.4	19.7
17.8	19.0	19.2	15.5	26.2	19.1	20.9	18.0	21.0
20.2	18.3	19.2	17.2	19.8	19.5	20.0	18.4	15.9
19.9	16.4	18.4	17.8	23.0	19.4	20.3	18.2	13.1
20.3	18.5	24.1	14.3	17.8	19.9	23.5	19.7	19.3
20.6	18.3	20.8	17.6	18.1	19.7	19.1	19.5	23.5
18.5	20.0	22.4	18.8	16.2	15.6	15.5	18.5	19.0

جدول (١ ـ ٣) تصفيف القياسات الواردة في الجدول (١ ـ ٢)

الرقمان الأول والثاني	الرقم الثالث	التعداد
12	2	1
13	1	1
14	3	1
15	5 6 5 9	4
16	8 4 2 1 9 2	6
17	4 8 8 3 2 8 6 8 8 1 8	11
18	5 5 4 3 5 3 4 8 3 1 4 0 4 2 5	15
19	9 7 0 3 2 2 4 5 8 1 5 4 9 7 7 1 7 5 7 3 0	21
20	2 3 6 0 8 9 5 6 8 9 0 3	12
21	7 6 4 1 0	5
22	7 4 0	3
23	3 2 0 5 0 5	6
24	1 2 8	3
25		0
26	2	1

ونلاحظ أن عملية التصفيف هذه هي، في الواقع، عملية فرز وتوزيع القياسات إلى فئات طول كل منها يساوي عشرة أمثال الواحد في المنزلة العشرية الأخيرة من قياسات البيان. أي أن طولها يساوي الواحد الصحيح إذا كانت القياسات معطاة لرقم عشري واحد وطولها واحد في العشرة، إذا كانت القياسات معطاة لرقمين عشريين،

وطولها عشر وحدات إذا كانت القياسات أعدادا صحيحة، وهكذا*.

وقد أصبح الجواب على التساؤل الذي طرحناه سهلا وميسورا، فنظرة إلى الجدول (١ ـ ٣) تبين أن ثلاثة عشر عاملا من بين التسعين عاملا، يقل مستوى الهيموغلوبين عندهم عن 17. وتكون النسبة المطلوبة $\frac{13}{90}$.

والجدير بالملاحظة أن كل ما خسرناه من المعلومات الواردة في البيان الأصلي (الخام) الوارد في الجدول (١ ــ ٢)، كنتيجة للتصفيف، هو الترتيب الزمني للحصول على القياسات. وقد لا يهمنا هذا في شيء أي أننا، عمليا، لم نخسر شيئا. ولكن وقفة تأمل هنا توضح لنا أن عملية التصفيف في بيانات تتضمن قياساتها أكثر من ثلاثة أرقام معنوية ستحتاج إلى جهود كبيرة، وكذلك ستكون الجهود كبيرة في حالة بيانات تتضمن عددا كبيرا من القياسات، مما يجعل التصفيف عملية غير رابحة في مثل تلك البيانات. فالجهود التي نبذلها في التصفيف قد لا تقل، بل قد تفوق، الجهود التي نحتاجها للإجابة على التساؤلات المطروحة مستخدمين البيان الأصلي مباشرة. وتبقى عملية التصفيف مقبولة فقط في بيانات من الحجم المتوسط، كالبيان المعطى في الجدول (١ ــ ٢) أو أصغر حجما، وفي مثل هذه البيانات، ونظرا لكبر عدد الفئات، تبقى إمكانية ظهور فئة خالية لا تتضمن أي قياس إمكانية قائمة، وهو أمر غير مستحسن.

وربما كان المثال السابق كافيا لتوضيح الفكرة التي نريد تقديمها، وهي أننا نحاول اختزال البيان الاحصائي الخام بطريقة تسمح لنا الإجابة عن تساؤلات، أو فهم نواح معينة مهمة من البيان الإحصائي، بسرعة وسهولة. وذلك لقاء فدية نقدمها، إذ نضحي ببعض المعلومات التي كان البيان الأصلي يوفرها لنا، ولكن البيان المختزل لم يعد قادرا على توفيرها. وسنقدم الآن اتجاها عاما ومفيدا لاختزال بيان إحصائي فيما يسمى بجداول التوزيع التكرارية.

---

* تسمى هذه الطريقة في التصفيف طريقة «الجذع والورقة»

مثال (١ ـ ٣)

قدمنا لخمسين مستجدا من طلبة الجامعة اختبارا لقياس «حاصل الذكاء» وكانت درجاتهم كما يلي:

جدول (١ ـ ٤). قياسات حاصل الذكاء لخمسين مستجداً

97	110	105	96	109	94	108	117	107	110	82	99	93
116	126	124	108	90	118	116	124	114	101	112	120	113
110	101	103	115	107	102	123	106	105	106	120	100	107
119	120	112	92	103	88	104	97	101	109	105		

إذا قمنا بتصفيف قياسات هذا البيان فسنجد الجدول (١ ـ ٥).

جدول (١ ـ ٥). تصفيف القياسات الواردة في الجدول (١ ـ ٤)

الرقمان الأول والثاني	الرقم الأخير	التعداد
08	2 8	2
09	3 9 4 6 7 0 7 2	8
10	7 8 9 5 1 8 7 0 6 5 6 2 7 3 1 5 9 1 4 3	20
11	0 7 0 3 2 4 6 8 6 5 0 2 9	13
12	0 4 4 6 0 3 0	7

توزعت القياسات على الفئات الخمس في الجدول (١ ـ ٥) فكان نصيب الفئة الأولى 2، وهي تتضمن جميع القياسات التي تنتمي إلى الفترة (90 ,80]، وكان نصيب الفئة الثانية 8، وهي تتضمن جميع قياسات البيان الإحصائي التي تنتمي إلى الفترة (100 ,90]. وكان نصيب الفئة الثالثة 20، وهي تتضمن جميع قياسات البيان التي تنتمي إلى الفترة (110 ,100]. وكان نصيب الفئة الرابعة 13، وهي تتضمن جميع قياسات البيان التي تنتمي إلى الفترة(120 ,110]. وكان نصيب الفئة الخامسة والأخيرة 7، وهي تتضمن جميع قياسات البيان الإحصائي التي تنتمي إلى الفترة (130 ,120].

بصورة عامة، لماذا لا نختار طول الفئة وبالتالي عدد الفئات بالشكل الذي نراه مناسبا للحالة المدروسة بدلا من أن تُفرض علينا كما هو الحال هنا؟ ولماذا لا نزيد من مقدار التضحية بمعلومات البيان الأصلي، ذات النفع البسيط للنواحي التي يتركز عليها اهتمامنا لقاء مزيد من توفير الجهود وسهولة العرض والحساب؟ فنحن مثلا قد لا نحتاج إلى الاحتفاظ بمفردات البيان الإحصائي، وإنما يقتصر اهتمامنا على معرفة كيفية توزعها على فئات نحددها سلفا تحديدا لا لبس فيه.

وإن أول ما تجدر معرفته هو مدى تغير القياسات في البيان الإحصائي. وباستعراض بسيط للقياسات نجد أن أصغر قياس هو 82، وأن أكبر قياس هو 126. ونقول إن مدى البيان الإحصائي هو الفرق بين أكبر قياس وأصغر قياس فيه، أي:

$$\text{المدى} = 126 - 82 = 44$$

لنقسم هذا المدى إلى عدد من الفئات نختاره بصورة كيفية. ولنأخذ هنا، مثلا، تسع فئات طول كل منها خمسة، فتكون الفئات كما يلي:

$$82 - 86, 87 - 91, 92 - 96, \ldots, 117 - 121, 122 - 126$$

ولنرتب جدولا مثل الجدول (١ ـ ٦). حيث نضع في العمود الأول حدود الفئات، وفي العمود الثاني، وسميناه عمود الفرز، نضع حذاء الفئة خطا مائلا في مقابل كل قياس في البيان ينتمي إلى هذه الفئة. ولسهولة التعداد تظهر كل حزمة من خمسة خطوط على حدة، ويقطع الخط الخامس الخطوط الأربعة السابقة له. ويسمى عدد القياسات التي تنتمي إلى الفئة i، مثلا، تكرار الفئة i، ونرمز له عادة بـ f_i. (f_1 تكرار الفئة الأولى، f_2 تكرار الفئة الثانية، . . .، وهكذا). وتظهر هذه التكرارات في العمود الثالث، وهي ناتجة عن تعداد الخطوط المقابلة للفئة في عمود الفرز. ونجد في العمود الرابع، التكرار النسبي، وهو يساوي التكرار مقسوما على العدد الكلي للقياسات $n = 50$. ونلاحظ أن مجموع عمود التكرار يجب أن يساوي 50، وأن مجموع عمود التكرار النسبي يجب أن

يساوي الواحد تماما. وإذا كان عمود التكرار يعطي عدد القياسات في البيان الإحصائي التي تنتمي إلى الفئة المقابلة فإن عمود التكرار النسبي هو تعبير آخر عن الفكرة نفسها، إذ يُقدِّم ذلك العدد على شكل نسبي (منسوبا إلى عدد القياسات الكلي) وسنلمس فيما بعد فائدة التعبير عن التكرار بالشكل النسبي.

جدول (١ ـ ٦). التوزيع التكراري لقياسات حاصل الذكاء لخمسين مستجدا

التكرار النسبي	التكرار f_i	الفرز	حدود الفئات
1/50	1	/	82 - 86
2/50	2	//	87 - 91
4/50	4	𝍸	92 - 96
7/50	7	𝍸 //	97 - 101
9/50	9	𝍸 𝍸	102 - 106
10/50	10	𝍸 ////	107 - 111
7/50	7	𝍸 //	112 - 116
6/50	6	//// /	117 - 121
4/50	4	////	122 - 126
1	50		المجموع

وترتيب القياسات كما في الجدول (١ ـ ٦) يسمى توزيعا تكراريا للقياسات. وبصورة عامة، التوزيع التكراري هو أي ترتيب للبيان الإحصائي يظهر توزّع قياساته على فترات معرفة بصورة اختيارية تسمى «فئات».

وقد اختيرت الفئات بصورة كيفية. ومن أجل البيان الإحصائي نفسه يمكن أن يختلف جدول التوزيع التكراري باختلاف تعريف الفئات وعددها، وليس هناك جدول يمكن القول إنه صحيح وما عداه من الجداول التي كان يمكن الوصول إليها غير

صحيحة، ولكن بعض هذه الجداول أفضل من بعض من حيث مقدرتها على تبيان النواحي المهمة في البيان الإحصائي دون الاحتفاظ بكثير من التفاصيل.

وبصورة عامة، يستحسن ألا يقل عدد الفئات عن خمس ولا يزيد على عشرين، تفادياً لظهور فئات خالية عند استكمال عملية الفرز، ذلك لأننا قد نخسر أكثر مما يجب من المعلومات إذا قل عدد الفئات عن خمس، وقد نحتفظ بما لا ضرورة له من التفاصيل عندما يزيد عدد الفئات على عشرين.

ويجب تعريف حدود الفئات بصورة واضحة لا تترك أي لبس في عملية الفرز، وتضمن انتماء كل قياس في البيان الإحصائي إلى فئة واحدة وواحدة فقط.

(١ ـ ٢) أنواع البيانات الإحصائية

تنقسم البيانات الإحصائية العددية إلى نوعين، أحدهما منفصل وتكون قياساته ناتجة عن عملية عد أو تعداد، مثل عدد حوادث المرور اليومي خلال فترة زمنية محددة، أو العدد السنوي لحالات الولادة، أو الزواج، أو الوفاة، أو الطلاق، في بلد معين، وتكون مثل هذه القياسات، دائما، أعدادا صحيحة. والنوع الآخر هو النوع المتصل (أو المستمر)، وتكون قياساته ناتجة عن استخدام جهاز أو أداة للقياس، مثل بيانات تتضمن قياسات طول، أو وزن، أو د رجة حرارة، أو مستوى التحصيل الدراسي، أو حاصل الذكاء، الخ.

وفي البيانات المستمرة، نفهم من العدد المقدم لنا شيئين، أولهما تصور عن مقدار الشيء المقيس، وثانيهما درجة الدقة التي سمح بها جهاز القياس المستخدم. والقول بأن طول شخص هو 167.5سم، يعطينا فكرة عن ارتفاع قامة الشخص، ويعطينا أيضا أن القياس جرى بدقة تصل إلى أقرب ملليمتر. أي أن آخر رقم معطى على اليمين، هو رقم مشكوك فيه. ولو أننا استخدمنا جهازا أكثر دقة، لحصلنا على قياس واقع في مكان ما بين 167.45 و 167.549. وأينما وقع هذا القياس فسيؤدي التدوير إلى الرقم العشري

الأول إلى العدد 167.5سم . وللسهولة جرت العادة على القول بأن عددا مثل 167.5سم يعني أي شيء بين 167.45 و 167.55سم .

ولأسباب عدة، نأخذ في الغالب، الحدود المضبوطة للفئة بعين الاعتبار، ونسميها «الحدود الحقيقية للفئة» أو «نهايتي الفئة». لنأخذ الفئة 86 - 82 في الجدول (١ ـ ٦)، فالقياس 82 يعني أي شيء بين 81.5 و 82.5، ويعني الـ 86، أي شيء بين 85.5 و 86.5. وهكذا يتراوح المدى الحقيقي للفئة بين 81.5 و 86.5. ويسمى هذان العددان «الحدان الحقيقيان للفئة» أو «نهايتا الفئة». وتجدر ملاحظة أن تطابق نهاية فئة مع بداية الفئة التي تليها، لا يؤدي إلى أي التباس في عملية الفرز، فالعدد 86.5 الذي يشكل حدا أعلى للفئة الأولى وحدا أدنى للفئة الثانية لا يمكن أن يظهر كقياس في البيان الاحصائي الأصلي ما دامت القياسات جميعها أعدادا صحيحة .

وهناك طرق أخرى يمكن استخدامها للتعبير عن حدود الفئة فمثلا يمكن، في المثال (١ ـ ٢ ـ ٣)، كتابة الفئات على الشكل :

82-, 87-, 92-, 97-, 102-, 107-, 112-, 117-, 122-

ونقصد بـ-82 جميع الأعداد الواقعة ضمن الفترة (87-82]، أي الأعداد بدءا من 82 إلى أقل من 87، وهكذا .

أو يمكن كتابتها على الشكل :

-87, -92, -97, -102, -107, -112, -117, -122, -127

ونقصد بـ 92- جميع الأعداد الواقعة ضمن الفترة (92-87]، أي الأعداد بدءا من 87 إلى أقل من 92. وهكذا .

وسنستخدم في هذا الكتاب الحدود الحقيقية للفئات في جميع البيانات سواء أكانت مستمرة أم منفصلة. واستخدامها في البيانات المنفصلة يضمن استمرارية

الأشكـال التي تمثل الجدول التكـراري بيانيـا كما سنـرى في الفقرة التـاليـة ، وهذا أمـر مستحسن . كما سنستفيد منه في أكثر من مكان في الفصول المقبلة .

ونلاحظ بوضوح أن طول الفئة مسـاو للفرق بين حديها الحقيقين . أما مركز الفئة فهـو منتصف المسافـة بين حديها ، ولحسـاب قيمته نأخـذ نصف مجموع الحدين ، ولـو استخدمنا الحدود(87-82] ، (92-87] الخ . فإن مركز الفئة الأولى سيكون 84.5 والثانية 89.5 الخ . وعند استخـدام الحدود الحقيقية [81.5-86.5] ، [86.5-91.5] الخ . فإن مركـز الفئة الأولى سيكون 84 والثانية 89 إلخ . واستخـدام الحدود الحقيقيـة إلى جانب أنـه يعـالج مشكلة وجود فراغات بين الفئات المتتاليـة ويضمن تطابق نهاية فئة مع بداية الفئة التي تليها بطريقة منطقية وعادلة ، فإنه يؤدي أيضا إلى حسابات أكثر دقة بصورة عامة .

وإذا أضفنا طول الفئـة إلى مركز الفئة الأولى حصلنـا على مركز الفئة الثـانية التي تليها وهكـذا . وسنتصور وجود فئـة على يسار الفئـة الأولى ، وبالطبع سيكون تكـرارها مساويا للصفر ولذلك سنسميها الفئة الصفرية على اليسار ، كما سنتصور وجود فئة على يمين الفئة الأخيرة ، وبما أن تكرارها صفر فسنسميها أيضا الفئة الصفرية على اليمين .

ولقد ذكرنـا أن تصنيف القياسات في فئات ، يهدف إلى تيسير عـرض البيانات ، وسهولـة القيام بحسـاب معايير إحصائيـة مفيدة في وصف وتحليل البيـان الإحصائي ، واستنتـاج معلومـات عامـة منه . وننطلـق في هذا من نـوعين من الافتراضات المتعلقـة بكيفية توزع القياسات ضمن الفئة الواحدة .

(i) عند حسـاب بعض المعايير الإحصائيـة ، أو عند استخـدام الطرق البيانيـة لعرض معلـومـات إحصـائية ، نفترض أن القيـاسات الـواقعة ضمن فئـة واحدة تتـوزع بانتظـام على الفترة الممتـدة بين نهـايتي الفئـة . وفي الجـدول (١ ـ ٦) ، مثـلا ، ينتمي عشر قياسات إلى الفئة 111-107 . والفترة الممتدة بين نهايتي الفئـة هي الفترة (106.5, 111.5) . ونفترض أن القيـاسات العشرة تتـوزع بـانتظـام فوق الـوحـدات الخمس التي تتألف منهـا الفئـة . أي نفترض ، كما يبين الجـدول (١ ـ ٧) ، قياسين بين 106.5 و 107.5 ، وقياسين بين 107.5 و 108.5 ، الخ .

جدول (١ ـ ٧). توزع القياسات بانتظام ضمن الفئة الواحدة

الفئة الجزئية	106.5-107.5	107.5-108.5	108.5-109.5	109.5-110.5	110.5-111.5
التكرار	2	2	2	2	2

(ii) والافتراض الثاني الذي نستخدمه عند حساب بعض المعايير الإحصائية هو اعتبار مركز كل فئة ممثلا لجميع القياسات التي تنتمي إليها، أي نفترض أن كل قياس من القياسات التي تنتمي إلى فئة مساوٍ لمركز الفئة.

والجدير بالذكر أن هناك بيانات إحصائية غير عددية تتضمن أصنافا معبرا عنها على شكل كلمات وصفية أو رموز، (علمي، أدبي)، (ذكر، أنثى)، (ممتاز، جيد جدا، جيد، مقبول، ضعيف) وبحيث ينتمي كل عنصر يخضع للتصنيف إلى صنف واحد منها فقط. ومثل هذه البيانات تسمى بيانات وصفية. وإذا أمكن تعريف ترتيب على هذه الأصناف أو الرموز يسمى البيان عندئذ بيانا ترتيبيا. فمثلا، يمكن القول أن «ممتاز» يمثل الصنف الأعلى يليه «جيد جدا» يليه «جيد» إلخ. مما يجعل أي بيان يتضمن تقديرات ممتاز. . . إلى ضعيف بيانا ترتيبيا. ونلاحظ أن ذلك غير ممكن في التصنيف (علمي، أدبي) أو (ذكر، أنثى).

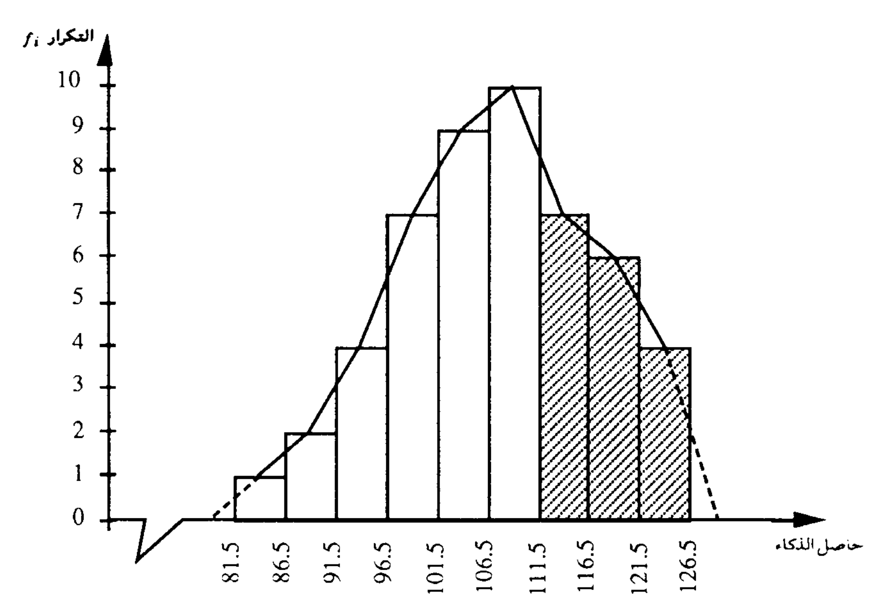
(١ ـ ٣) التمثيل البياني لتوزيع تكراري

يقدم لنا التمثيل البياني للمعلومات الإحصائية عونا كبيرا، فهو يسمح بإدراك الخواص الأساسية للتوزيع التكراري، ومقارنة توزيع تكراري بآخر. والتمثيل البياني هو صورة هندسية لجملة القياسات. وفي العديد من الحالات يسهّل رد مجموعة المعلومات الرقمية إلى صورة هندسية، فهم طبيعة المسألة الإحصائية، واستنباط الحلول المناسبة لها. وقد أصبح التمثيل البياني ممارسة شبه يومية في حياتنا. فالصحف والمجلات، والنشرات التجارية، وتقارير الأعمال والمشاريع، والدوريات العلمية المختلفة،

والتقارير الحكومية، تستخدم جميعها، وعلى نطاق واسع، التمثيل البياني. وهناك تفرعات كثيرة لوسائل التعبير البياني عن جملة من المعلومات الإحصائية، وسنقتصر هنا على ذكر أكثرها أهمية وفائدة في مجالات الاستقراء الإحصائي. ويمكن لمن أراد الاستزادة العودة إلى بعض المراجع المذكورة في نهاية الكتاب.

(١ ـ ٣ ـ ١) المدرج التكراري

لرسم المدرج التكراري، نتخذ المحور الإحداثي السيني لتمثيل الفئات، ونحدد عليه النقاط التي تمثل نهايات الفئات (حدودها الحقيقية). ونتخذ المحور الإحداثي الصادي لتمثيل التكرار f_i. ثم نرسم فوق الفترة الممتدة بين نهايتي كل فئة مستطيلا يرتفع بمقدار التكرار المقابل لهذه الفئة. ونجد في الشكل (١ ـ ١) المدرج التكراري الموافق للتوزيع التكراري المعطى في الجدول (١ ـ ٦).



شكل (١ـ١). مدرج التكرار ومضلع التكرار لقياسات حاصل الذكاء لخمسين مستجدا

وتمثل المساحة تحت مدرج التكرار نوعا من أنواع التمثيل الهندسي لجملة القياسات في البيان الإحصائي. فإذا نظرنا إلى تكرار كل فئة بأنه مساهمة الفئة في تركيبة

البيان الإحصائي، إذا جاز التعبير، فإن مساحة المستطيل المقام فوق الفئة يتناسب مع هذه المساحة. وكلما كان التكرار أكبر ارتفع المستطيل وزادت مساحته. ولو تساءلنا في المثال (١ ـ ٣) عن نسبة الطلبة الذين نالوا درجات أعلى من 111.5 لوجدنا أن هذه النسبة تساوي نسبة المساحة تحت مدرج التكرار الواقعة على اليمين من 111.5 (وهي مساحة المستطيلات الثلاثة الأخيرة المظللة في الشكل (١ ـ ١) إلى المساحة الكلية تحت مدرج التكرار. وما دامت الفئات جميعها بالطول نفسه، أي ما دامت قواعد المستطيلات المرسومة متساوية وتبقى ثابتة من فئة إلى أخرى، فإن مساحة كل مستطيل تتناسب مع ارتفاعه (مع تكرار الفئة) ونسبة المساحة المظللة إلى المساحة الكلية هي في الواقع نسبة مجموع التكرارات الموافقة للفئات الثلاث الأخيرة إلى العدد الكلي للقياسات، أي $\frac{17}{50}$ أو 34% وهي النسبة المطلوبة بالضبط.

ويمكن اللجوء إلى هذا المبدأ في تمثيل المدرج التكراري سواء أكانت أطوال الفئات متساوية أم لا. ففي حالة رسم مدرج تكراري لتوزيع تكراري لا تتساوى فيه أطوال الفئات، إما كنتيجة لطبيعة التفاصيل التي رؤي أن يحتفظ بها الجدول التكراري، أو نتيجة لدمج عدة فئات، تكراراتها صغيرة نسبيا، بعضها مع بعض لتشكل فئة واحدة؛ لا بد من القيام بتعديلات مناسبة تأخذ في الاعتبار أطوال الفئات، وتجعل المساحة المقامة فوق فئة، متناسبة مع تكرار هذه الفئة. ويكون رسم مستطيلات ارتفاعاتها مساوية لتكرار الفئة غير صحيح. وبذلك نتجنب رسم مدرج تكرار يعطي انطباعات مضللة إلى حد بعيد. ونوضح الفكرة وطريقة العمل من خلال المثال التالي.

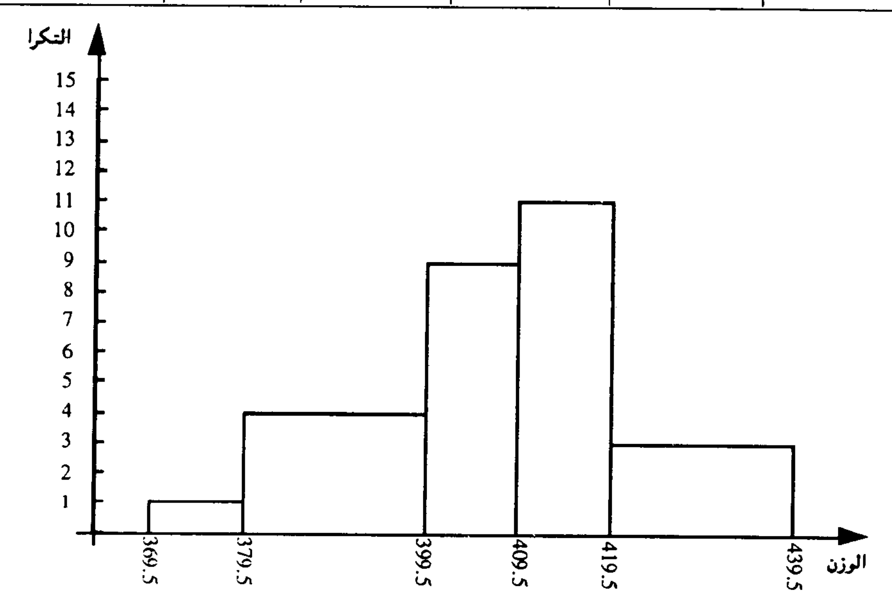
مثال (١ ـ ٤)

فيما يلي جدول توزيع تكراري لأوزان 35 فأرا مقاسة إلى أقرب غرام. ارسم المدرج التكراري.

الرسم مبين في الشكل (١ ـ ٢) حيث عدّلنا في ارتفاع المستطيل المقام فوق كل فئة بحيث نحفظ تناسب المساحة المرسومة فوق الفئة مع التكرار الموافق، والفئة الثانية والخامسة لها أطوال مضاعفة ولذلك رسمنا فوق كل منها مستطيلا ارتفاعه يساوي

جدول (١ ـ ٨): التوزيع التكراري لأوزان35 فأرا

حدود الفئات	370-379	380-399	400-409	410-419	420-439
التكرار	1	8	9	11	6



شكل (١ ـ ٢). المدرج التكراري لأوزان 35 فأرا

نصف التكرار الموافق للفئة (4 في الفئة الثانية و 3 في الفئة الخامسة). إن المساحة الكلية للمدرج هي:

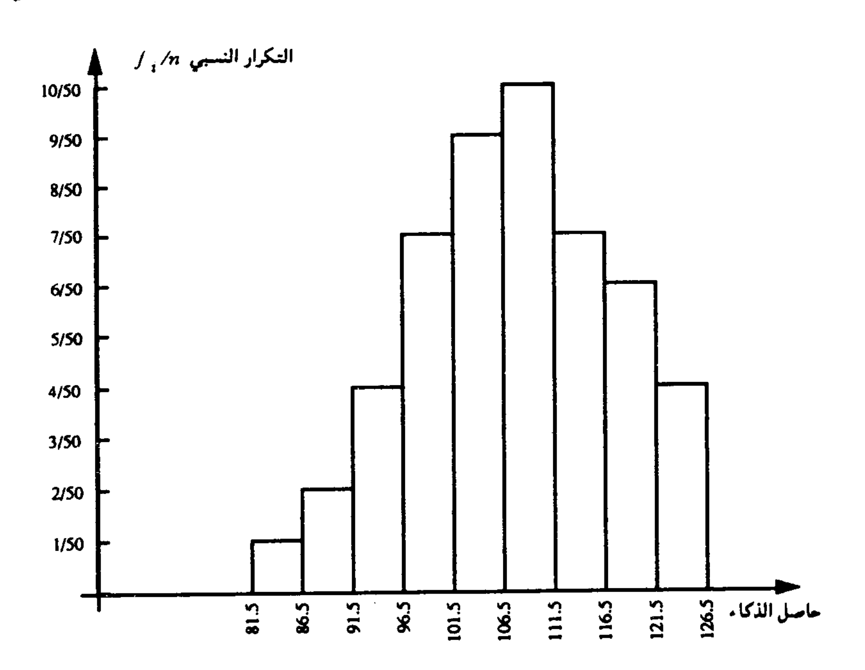
$$10 \times 1 + 20 \times 4 + 10 \times 9 + 10 \times 11 + 20 \times 3 = 350$$

ونسبة مساحة المستطيل الموافق لكل فئة إلى المساحة الكلية تساوي تماما نسبة تكرار الفئة إلى مجموع التكرارات. فمثلا نسبة مساحة المستطيل الثاني إلى المساحة الكلية هي $\frac{80}{350}$ وهي تساوي $\frac{8}{35}$.

(١ ـ ٣ ـ ٢) مدرج التكرار النسبي

لا تختلف طريقة رسم مدرج التكرار النسبي عن مدرج التكرار سوى أن المستطيل الموافق لكل فئة يرتفع الآن بما يساوي التكرار النسبي للفئة. ولكي نحافظ على ارتفاع ووضوح مناسبين للصورة الناتجة، لا بد أن تكون وحدة الطول على المحور الصادي أكبر بصورة مناسبة مما كانت عليه على المحور الصادي لمدرج التكرار. ولو كان لدينا n قياسا، وكبرنا وحدة الطول على المحور الصادي n مرة، لحصلنا على صورة لمدرج التكرار النسبي مطابقة تماما لصورة مدرج التكرار. وكل ما في الأمر أن التدريج 1 على المحور الرأسي أصبح الآن $\frac{1}{n}$ ، والتدريج 2 أصبح $\frac{2}{n}$ ، وهكذا. ونجد في الشكل (١ ـ ٣) مدرج التكرار النسبي لقياسات حاصل الذكاء لخمسين مستجدا.

تجدر ملاحظة أنه ليس من الضروري، عند رسم شكل بياني، أن تكون وحدة الأطوال نفسها على المحورين. ونتخذ وحدة الطول على كل من المحورين لتشغل الصورة الناتجة الحيز المخصص لها، وتتخذ موقعا مناسبا في الاتجاهين الأفقي والرأسي،



شكل (١ ـ ٣). مدرج التكرار النسبي لقياسات حاصل الذكاء لخمسين مستجدا

تماما كما تتخذ الصورة المخرجة في التليفزيون موقعها على الشاشة المخصصة لها، فلا هي منحازة إلى يمين الشاشة ولا إلى يسارها، ولا هي مرتفعة أو منخفضة أكثر مما ينبغي. وتكبير وحدة الطول على المحور السيني يؤدي إلى توسيع الصورة في الاتجاه الأفقي يمينا ويسارا. وتكبير وحدة الطول على المحور الصادي يؤدي إلى تمدد الصورة في الاتجاه الرأسي علوا وهبوطا. وأفضل ترتيب لوحدتي الطول هاتين، هو ذلك الذي يكفل وضوح الصورة، ويُخرجها بحيث تشغل الحيز المخصص لها بشكل مناسب.

ويبقى هذا كله صحيحا في حالة فئات غير متساوية أيضا، فصورة مدرج التكرار النسبي تتطابق مع صورة مدرج التكرار عندما نجعل التدريج 1 على المحور الصادي مساويا $1/n$، والتدريج 2 مساويا لـ$2/n$، وهكذا. والشكل (١ ـ ٢)، يصبح صورة لمدرج التكرار النسبي للتوزيع التكراري لأوزان 35 فأرا المعطى في الجدول (١ ـ ٨)، إذا اعتبرنا التدريج 1 على المحور الصادي مساويا الآن لـ 1/35 والتدريج 2 مساويا لـ 2/35، إلخ.

ومع أن اهتمامنا المباشر، في المثال (١ ـ ٣)، ينصب على وصف القياسات الخمسين، إلا أننا نهتم أكثر بالمجتمع الذي أخذنا منه هذه القياسات. ويمكن النظر إلى القياسات الخمسين كعينة مأخوذة من مجتمع طلبة السنة الأولى في جامعة أو عدد من الجامعات. وفي جميع الأحوال، لو توفرت لنا قياسات حاصل الذكاء لعناصر المجتمع كلها، لأمكن، بالطريقة ذاتها، إقامة المدرج التكراري للمجتمع.

لندرس الآن مدرج التكرار النسبي في الشكل (١ ـ ٣) بتفصيل أكثر. فلو افترضنا أن طول الفئة (وهي تساوي خمس وحدات) أصبحت وحدة قياس جديدة، أي أن طول الفئة بالوحدة الجديدة هو الواحد، فستصبح مساحة المستطيل المقام فوق الفئة مساوية للتكرار النسبي الموافق لهذه الفئة، وستصبح المساحة الكلية تحت مدرج التكرار النسبي مساوية للواحد تماما. ولنسأل الآن، ما هي نسبة الطلاب الذين يزيد حاصل ذكائهم على 111.5 مثلا؟ بالعودة إلى مدرج التكرار النسبي نرى أن هذه النسبة تشمل كل الفئات على اليمين من 111.5 . وبالاستفادة من الجدول (١ ــ ٦) نرى أن

سبعة عشر مستجدا حصلوا على أكثر من 111.5. أي أن النسبة المطلوبة هي 17/50 أو 34%. ونلاحظ أن هذه النسبة هي أيضا المساحة تحت مدرج التكرار النسبي في الشكل (١ ـ ٣) التي تقع على يمين 111.5 .

(١ ـ ٣ ـ ٣) مضلع التكرار

نأخذ منتصفات القواعد العليا للمستطيلات في مدرج التكرار، ونصل بينها بخطوط مستقيمة، فنحصل على ما يسمى بمضلع التكرار، أي أننا لو حددنا من أجل كل فئة نقطة إحداثيها السيني هو مركز الفئة، وإحداثيها الصادي هو تكرار الفئة، ثم وصلنا بين هذه النقاط بقطع مستقيمة لحصلنا على مضلع التكرار. ويمكن رسم مدرج التكرار ومضلع التكرار على الشكل نفسه، أو في شكلين منفصلين. ونجد في الشكل (١ ـ ١) مضلع التكرار للتوزيع التكراري في الجدول (١ ـ ٦) .

ويمكن إغلاق مضلع التكرار على الجانبين بوصل أول نقطة منه بمركز الفئة الصفرية على اليسار، ووصل آخر نقطة منه بمركز الفئة الصفرية على اليمين. (انظر الشكل (١ ـ ١) حيث رسمنا هاتين الوصلتين بخط منقط.)

ونلاحظ وجود فرق بسيط بين المساحة تحت مضلع التكرار والمساحة تحت مدرج التكرار ويتناقص هذا الفرق كلما ازداد عدد الفئات وصغر طول الفئة. هذا بصورة عامة، أما إذا كانت أطوال الفئات متساوية فالمساحتان متساويتان.

(١ ـ ٤) مضلع التكرار المتجمع الصاعد

من الخصائص المهمة للبيان الإحصائي معرفة العدد الذي تقل عنه نسبة معينة من القياسات، أو معرفة النسبة من القياسات التي تقل عن قيمة معينة، أو نسبة القياسات التي تتجاوز قيمة معينة. ففي بيان من الدرجات في مسابقة عامة، يمكن أن نعتبر العدد الذي يقل عنه تسعون بالمائة من القياسات الحد الفاصل بين تقدير الممتاز وما دون الممتاز. وفي بيان يمثل مستويات الهيموغلوبين في الدم تمثل نسبة القياسات، التي تقل عن قيمة معينة، نسبة المصابين بفقر الدم. وفي بيان يمثل

معدلات التوتر الشرياني (ضغط الدم) تمثل نسبة القياسات التي تزيد على قيمة معينة نسبة المصابين بمرض فرط التوتر الشرياني (ارتفاع معدل ضغط الدم).

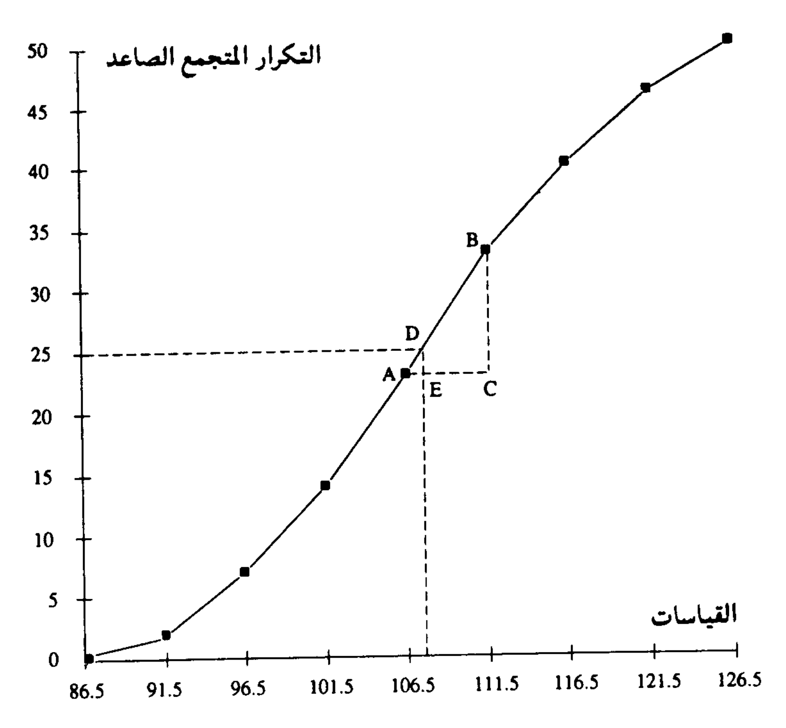
لنعد إلى الجدول (١ ـ ٦) فما هو العدد الذي يقل عنه خمسون بالمائة من القياسات؟ أو ما هي النقطة التي يقع إلى اليسار منها خمس وعشرون بالمائة من القياسات؟ أو ما هي نسبة القياسات التي تقل عن 101.5؟ إلخ.

وللجواب على مثل هذه التساؤلات، بصورة تقريبية وسريعة، نقيم جدول التكرار المتجمع الصاعد كما في الجدول (١ ـ ٩)، حيث نضع في العمود الأول، وعنوانه «أقل من»، الحدود الحقيقية العليا للفئات، ونضع في العمود الثاني، وعنوانه «التكرار المتجمع الصاعد» عدد القياسات الموافق.

جدول (١ ـ ٩). جدول التكرار المتجمع الصاعد لقياسات حاصل الذكاء لخمسين مستجدا

التكرار المتجمع الصاعد	أقل من
1	86.5
3	91.5
7	96.5
14	101.5
23	106.5
33	111.5
40	116.5
46	121.5
50	126.5

ولتمثيل الجدول بيانيا نعتمد المحور السيني محورا للقياسات، والمحور الصادي محورا للتكرار المتجمع الصاعد. ونرسم لكل فئة نقطة في مستوى الإحداثيات، إحداثيها السيني هو الحد الأعلى الحقيقي للفئة، وإحداثيها الصادي هو التكرار المتجمع الصاعد المقابل. ثم نصل بين النقاط الناتجة المتتالية بقطع مستقيمة فنحصل على مضلع يدعى «مضلع التكرار المتجمع الصاعد». (انظر الشكل (١ ـ ٤)).



شكل (١ ـ ٤). مضلع التكرار المتجمع الصاعد لقياسات حاصل الذكاء لخمسين مستجدا

وبالطريقة نفسها يمكن رسم مضلع التكرار النسبي المتجمع الصاعد وذلك باستخدام المحور الصادي لتمثيل التكرارات النسبية المتجمعة. ولو أن العمود الأول في الجدول (١ ـ ٩) تضمن الحدود الحقيقية الدنيا للفئات وكان عنوانه «أكثر من» لحصلنا على جدول تكرار متجمع نازل، ورسمه البياني بالطريقة السالفة ذاتها سيعطي مضلع التكرار المتجمع النازل. وسنترك ذلك تمرينا للطالب.

ولإيجاد القياس الذي يقع على اليسار منه 50% من القياسات، نحسب أولا رتبة القياس المطلوب $\frac{50}{100} \times n$ ، حيث n عدد القياسات، فنجد:

$$25 = \frac{50}{100} \times 50 = \text{رتبة القياس المطلوب}$$

أي أن القياس المطلوب ينتمي إلى الفترة [106.5, 111.5].

وسنحسب القياس المطلوب مفترضين أن القياسات التي تنتمي إلى فئة تتوزع بانتظام فوق الفترة التي تمتد بين نهايتي الفئة، أو بعبارة أعم مفترضين أن العلاقة بين القياس والتكرار فوق الفترة [106.5, 111.5] هي علاقة خطية تتمثل في معادلة المستقيم الواصل بين النقطتين A و B.

من تشابه المثلثين ADE و ABC نجد:

$$\frac{AE}{AC} = \frac{ED}{CB}$$

$$\frac{AE}{5} = \frac{2}{10}$$

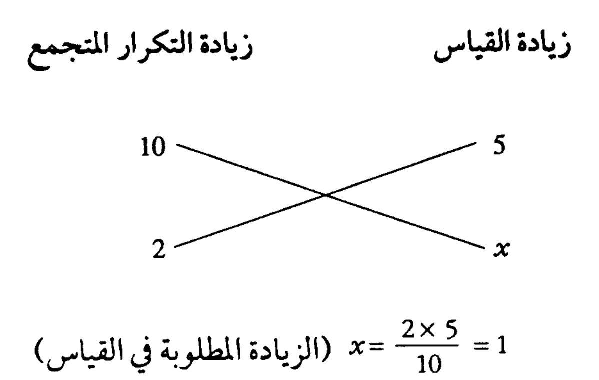
$$AE = \frac{5 \times 2}{10} = 1$$

ويكون القياس المطلوب، وهو الاحداثي السيني للنقطة D، مساويا الإحداثي السيني لـ A مضافا إليه AE، وهكذا نجد:

107.5 = 106.5 + 1 = القياس المطلوب

ويمكن القيام بهذه الحسابات معتمدين على جدول التكرار المتجمع الصاعد، ودون الحاجة إلى رسم مضلع التكرار، حيث نتبع المحاكمة التالية:

نرى من جدول التكرار المتجمع الصاعد أن 23 قياسا من القياسات الخمسين أقل من 106.5، وأن 33 قياسا أقل من 111.5. وبتطبيق التناسب الطردي نقول إنه عندما زاد التكرار المتجمع بمقدار 10، (من 23 إلى 33) زاد القياس بمقدار 5، (من 106.5 إلى 111.5). فما هي قيمة الزيادة في القياس عندما يزداد التكرار المتجمع بمقدار 2 فقط (من 23 إلى 25)؟



$x = \frac{2 \times 5}{10} = 1$ (الزيادة المطلوبة في القياس)

ويكون القياس المطلوب :

$$106.5 + 1 = 107.5$$

وإذا توفر ورق ميلليمتري نرسم عليه مضلع التكرار المتجمع الصاعد، فيمكن استخدام الرسم البياني لإيجاد القياس المطلوب، وهذا القياس ليس إلا الإحداثي السيني لنقطة على مضلع التكرار المتجمع الصاعد إحداثيها الصادي 25 . ولذلك نرسم من النقطة 25 على المحور الصادي خطا أفقيا يقطع مضلع التكرار المتجمع الصاعد في نقطة ننزل منها عمودا على المحور السيني فيقطعه في النقطة المطلوبة، وهي على الشكل (١ ـ ٤) حوالي 107.5 .

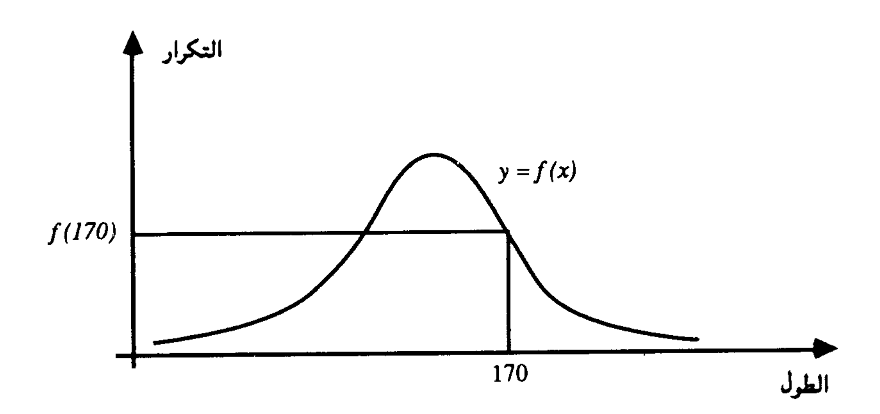
وسنجد فيما بعد أن هذا القياس يسمى الوسيط . وقد لخصنا هنا الطريقتين الحسابية والبيانية للحصول على الوسيط . ومن الواضح أنه يمكن تطبيق الطريقة نفسها لحساب القياس الذي يقع على اليسار منه 25% من القياسات، وبصورة عامة القياس الذي يقع على اليسار منه p بالمائة من القياسات، حيث p أي عدد بين الصفر والمائة، ويسمى مثل هذا القياس المئين p.

ولمعرفة نسبة القياسات التي تقل عن 101.5، مثلا، نرفع من النقطة 101.5 على المحور السني عمودا يقطع مضلع التكرار المتجمع في نقطة نرسم منها موازيا للمحور السيني فيقطع المحور الصادي في النقطة 14، وتكون النسبة المطلوبة $\frac{14}{50} = 28\%$.

(١ ـ ٥) منحنى التكرار

لنعد إلى مضلع الكرار في الفقرة (١ ـ ٣ ـ ٣). ولنفترض أننا صغرنا طول الفئة إلى نصف ما هو عليه. أي ضاعفنا عدد الفئات، ثم رسمنا مضلعا تكراريا، فسيتضاعف عندئذ عدد رؤوس هذا المضلع، وستقترب رؤوس المضلع بعضها من بعض. ولكن العدد البسيط من القياسات لا يسمح لنا بالمضي في مثل هذه العملية، لأنه قد يترك العديد من الفئات خالية وتكرارها صفر، مما يصيب المضلع بانقطاعات في أكثر من مكان، الأمر الذي لا يقلق كثيرا عندما يصف المضلع التكراري «مجتمعا» يتضمن عددا هائلا من القياسات. فلنتصور إذا، أن لدينا معينا لا ينضب من القياسات، أي لنتصور ظرفا يمكننا معه جعل طول الفئة أصغر فأصغر، وفي الوقت ذاته، زيادة عدد القياسات التي تخضع للتصنيف لتصبح أكبر فأكبر، ولندفع الآن مثل هذا التصور إلى نهاياته القصوى ليصبح طول الفئة صغيرا بلا حدود، ويصبح معه عدد القياسات الكلي كبيرا بلا حدود، فسنصل عندئذ إلى خط ناعم مستمر، لا انكسارات فيه ولا زوايا، يسمى منحنى التكرار. وعندئذ يقابل كل قياس على المحور السيني إحداثي صادي يتناسب مع تواتر ظهور هذا القياس في المجتمع الذي يصفه منحنى التكرار.

ولنفرض، على سبيل المثال، أن منحنى التكرار في الشكل (١ ـ ٥) يصف ظاهرة توزع الطول في مجتمع من الذكور البالغين يتضمن عشرات الملايين، فالإحداثي الصادي للنقطة 170سم، مثلا، يمثل أو يتناسب مع تواتر ظهور الطول 170سم في هذا المجتمع. وبصورة عامة، نعتمد منحنيات التكرار كنماذج رياضية (نظرية) لتمثيل ظواهر عامة في حياتنا العملية. وعلى سبيل المثال، سنعرض فيما يلي إحصائيات (Kendall and Stuart, 1977) لثلاث ظواهر مختلفة تتناول عددا كبيرا من الأفراد. وسنجد أن مضلع التكرار لكل ظاهرة يوحي بشكل معين لمنحنى التكرار (أو النموذج) الذي يمكن اعتماده لوصف هذه الظاهرة. وسنتعرف في الفصل الرابع وما بعده على ما نقصده بكلمة «نموذج»، والدور الذي تلعبه النماذج في التطبيقات العملية للإحصاء.



شكل (١ ـ ٥). منحنى التكرار لتوزيع الطول في مجتمع من الذكور البالغين

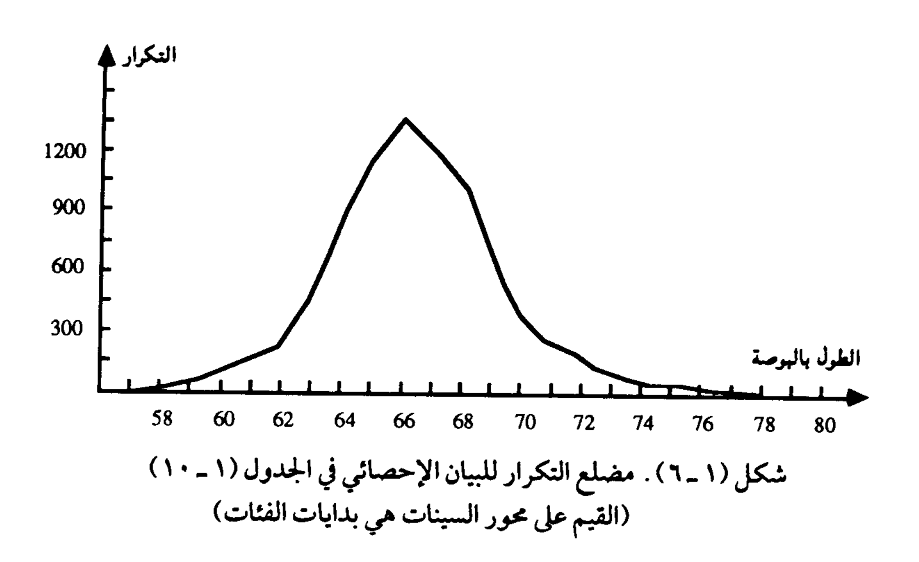
يبين الجدول (١ ـ ١٠) توزيع التكرار لأطوال 8585 ذكرا بالغا ممن ولدوا في المملكة المتحدة. وباعتبار أن دقة القياس كانت إلى أقرب $\frac{1}{8}$ من البوصة، فالحدود الحقيقية للفئات هي من $57\frac{15}{16} - 56\frac{15}{16}$ ، $58\frac{15}{16} - 57\frac{15}{16}$ ، وهكذا . . .

جدول (١ ـ ١٠). التوزيع التكراري لـ 8585 ذكرا بالغا ممن ولدوا في المملكة المتحدة

التكرار	الطول (بدون حذاء)	التكرار	الطول (بدون حذاء)
1230	68 -	2	57 -
1063	69 -	4	58 -
646	70 -	14	59 -
392	71 -	41	60 -
202	72 -	83	61 -
79	73 -	169	62 -
32	74 -	394	63 -
16	75 -	669	64 -
5	76 -	990	65 -
2	77 -	1223	66 -
		1329	67 -

8585 = مجموع التكرارات

وفي الشكـل (١ ـ ٦) نجد مضلــع التكرار، ومن الواضح أن هذا المضلع يقترح بقوة أن نموذجا على شكل الجرس (انظر الشكل (١ ـ ٥)) هو النموذج المناسب لتمثيل ظاهرة توزع الطول في مجتمع من الذكور البالغين في بيئة معينة .



شكل (١ ـ ٦) . مضلع التكرار للبيان الإحصائي في الجدول (١ ـ ١٠)
(القيم على محور السينات هي بدايات الفئات)

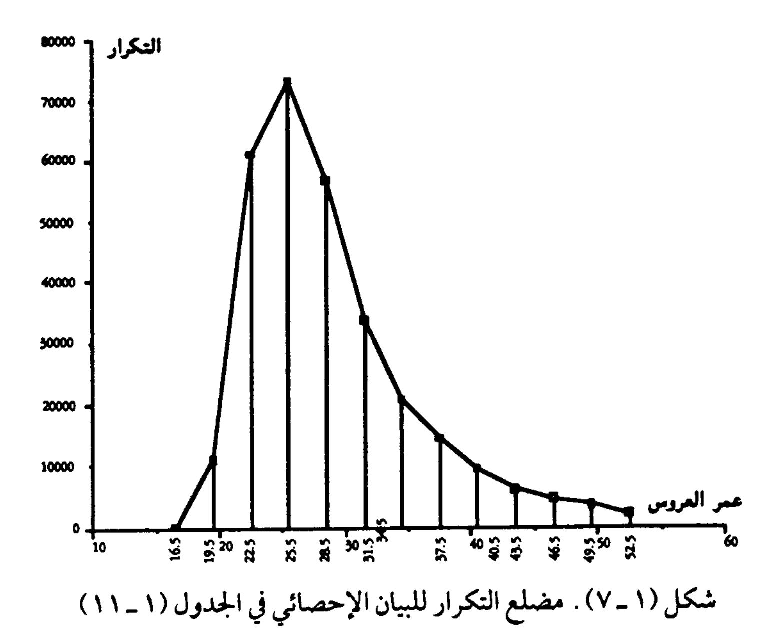
يبين الجدول (١ ـ ١١) توزيع التكرار لـ 301785 عقد زواج في استراليا بين 1907 و 1914، مصنفة وفقا لعمر العروس في فئات طول كل منها 3 سنوات .

والمضلع التكراري يقترح بـوضوح نموذجا يعـرف بنموذج «جامـا» وهو منحنى تكرار غير متماثل يتـزايد بسرعـة إلى قمة ثم ينحـدر منها بسرعة (سرعـة التزايـد وسرعة الانحدار تختلف من حالة إلى أخـرى) ليتهادى بعد ذلك متناقصا باطـراد تناقصا بطيئا مقتربـا من محور السينـات . ونقـول عن نمـوذج كهذا أنـه ملتـو إلى اليمين أو مـوجب الالتواء [انظــر الشكل (١ ـ ٧)]. ونجـد في الشكل (١ ـ ٨) منحنى تكـرار من النوع «جاما» .

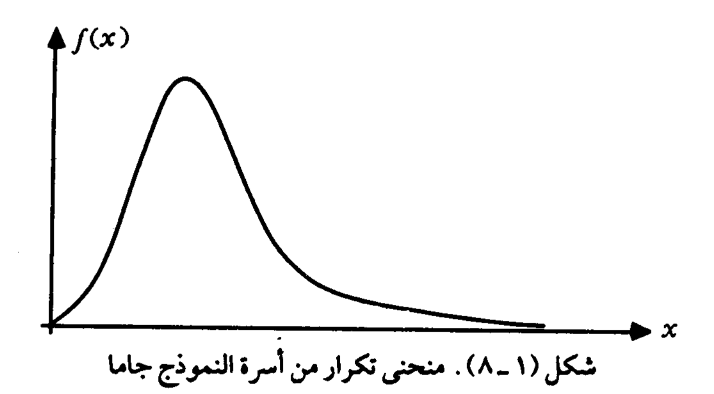
جدول (١ ـ ١١). التوزيع التكراري لـ 301785 عقد زواج في استراليا مصنفة وفق عمر العروس.

مركز الفئات	التكرار	مركز الفئات	التكرار
16.5	294	55.5	1655
19.5	10995	58.5	1100
22.5	61001	61.5	810
25.5	73054	64.5	649
28.5	56501	67.5	487
31.5	33478	70.5	326
34.5	20569	73.5	211
37.5	14281	76.5	119
40.5	9320	79.5	73
43.5	6236	82.5	27
46.5	4770	85.5	14
49.5	3620	88.5	5
52.5	2190		

301785 = مجموع التكرارات



شكل (١ ـ ٧). مضلع التكرار للبيان الإحصائي في الجدول (١ ـ ١١)



شكل (١ ـ ٨). منحنى تكرار من أسرة النموذج جاما

تمـريـــن

ارسم منحنيا ملتويا إلى اليسار.

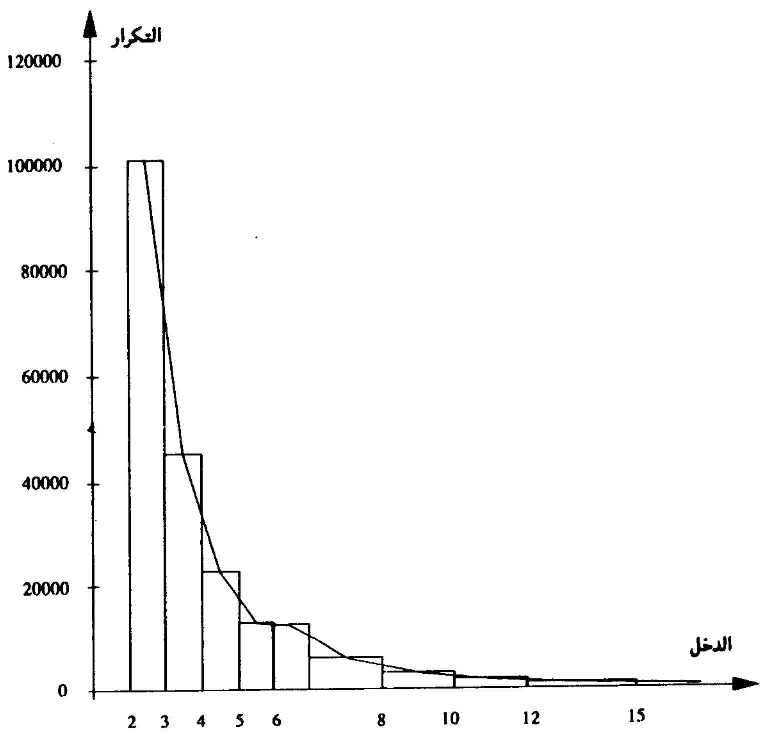
يبين الجدول (١ ـ ١٢) توزيع التكرار لـ 213938 شخصا في المملكة المتحدة مصنفين وفقا لشرائح الدخل مقدرة بآلاف الجنيهات.

جدول (١ ـ ١٢). توزيع التكرار وفق فئات الدخل لـ 213938 شخصا في المملكة المتحدة

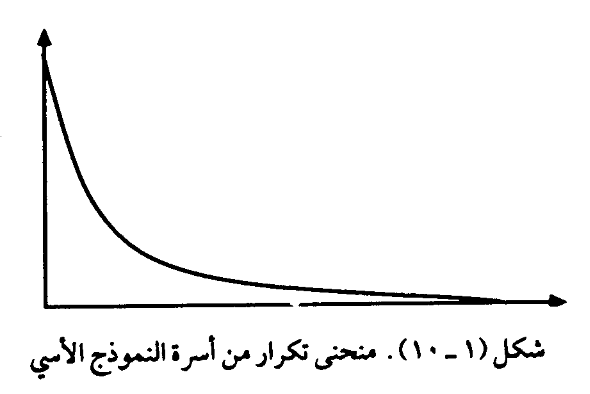
التكرار (عدد الأشخاص)	فئات الدخل بآلاف الجنيهات
101369	2 -
45532	3 -
23263	4 -
13475	5 -
13456	6 -
6419	8 -
3551	10 -
2926	12 -
2007	15 -
820	20 -
399	25 -
376	30 -
134	40 -
128	50 -
45	75 -
38	100 -

المجموع = 213938

ويقترح مضلع التكرار منحنى تكرار مناسب لهذه الظاهرة (توزع فئات الدخل في المملكة المتحدة) من النوع J. وتسمى هذه الأسرة من النماذج بأسرة النماذج الأسية. وهي تبدأ بقمتها ثم تنحدر بسرعة متقاربة إلى محور السينات. ونجد في الشكل (١ ـ ١٠) منحنى تكرار من أسرة النموذج الأسي.



شكل (١ ـ ٩). مدرج التكرار ومضلع التكرار للبيان الإحصائي في الجدول (١ ـ ٨)



شكل (١ ـ ١٠). منحنى تكرار من أسرة النموذج الأسي

تمارين (١ ـ ١)

١) تتغير أوزان خمسين طالبا مقاسة إلى أقرب «باوند» من 177 إلى 265. إذا أردت تصنيف هذه الأوزان في عشر فئات فاكتب حدود الفئات، والحدود الحقيقية للفئات؛ ومراكز الفئات. ما طول الفئة؟

٢) كانت مراكز الفئات لتوزيع تكراري لمجموعة من قياسات درجة الحرارة مأخوذة إلى أقرب درجة مئوية، كما يلي:

16, 25, 34, 43, 52, 61

أوجد:

أ ـ حدود الفئات؛ ب ـ الحدود الحقيقية للفئات.

٣) فيما يلي عدد الأميال التي قطعتها كل من أربعين سيارة إسعاف بجالون واحد من البنزين:

24.5	23.6	24.1	25.0	22.9	24.7	23.8	25.2	24.9
24.1	23.7	24.4	24.7	23.9	25.1	24.6	23.3	24.3
24.8	22.8	24.6	23.9	24.1	24.4	24.5	25.7	23.6
24.0	24.7	23.1	23.9	24.2	24.7	24.9	25.0	24.8
24.5	23.4	24.6	25.3					

أ ـ لخص هذا البيان الإحصائي في جدول توزيع تكراري متخذا الفئات:

22.5 - 22.9; 23.0 - 23.4, . . ., 25.5 - 25.9

ب ـ ارسم مدرج التكرار، ومدرج التكرار النسبي، ومضلع التكرار.

جـ ـ اكتب جدول التكرار المتجمع الصاعد.

د ـ ارسم مضلع التكرار المتجمع الصاعد

هـ ـ ما عدد القياسات التي هي أقل من 23.75؟، أكثر من 23.45؟، أقل من 24.3؟، وأقل من 25.2؟

و ـ ما القياس الذي يقل عنه خمسون بالمئة من القياسات؟ خمس وعشرون بالمئة من القياسات؟ وخمس وسبعون بالمئة من القياسات؟

٤) فيما يلي درجات 40 طالبا في اختبار ١٠٦ إحص:

42	88	37	75	98	93	73	62	96	80
52	76	66	54	73	69	83	62	53	79
69	56	81	75	52	65	49	80	67	59
88	80	44	71	72	87	91	82	89	79

أ ـ اكتب جدول التوزيع التكراري لهذا البيان الإحصائي مستخدما الفئات:

35 - 39; 40 - 44; . . .

ب ـ ارسم مدرج التكرار، ومدرج التكرار النسبي، ومضلع التكرار، مستخدما ورقة بيانية.

جـ ـ اكتب جدول التكرار المتجمع الصاعد وارسم مضلعه.

د ـ ما القياس المقابل لتكرار متجمع يساوي العشرين بالطريقتين الحسابية والبيانية؟

*٥) يتولى الإشراف الصحي على عدد من مدارس تعليم البنات 44 وحدة صحية منتشرة في أنحاء المملكة. وفيما يلي عدد المدارس المرتبطة بكل من هذه الوحدات الصحية (لا يتضمن البيان مدارس الرياض وجدة والإحساء ومكة المكرمة):

23,	46,	20,	30,	28,	12,	35,	50,	33,	65,	85,
24,	40,	50,	23,	40,	30,	50,	23,	20,	38,	68,
58,	15,	15,	100,	105,	6,	59,	36,	22,	89,	21,
35,	100,	42,	38,	58,	32,	62,	48,	32,	19,	56

* مأخوذ من التقرير الصحي السنوي الصادر عن وزارة الصحة لعام ١٤٠٦هـ، ص ٢٨٧.

متخذا الفئات 105 - 90 ,89 - 78 ,. . . ,41 - 30 ,29 - 18 ,17 - 6

أ ـ ارسم مدرج التكرار النسبي.
ب ـ ارسم مضلع التكرار المتجمع الصاعد
جـ ـ أوجد حسابيا وبيانيا المئين 10 والمئين 90.

٦) عند تلخيص بيان إحصائي حصلنا على التوزيع التكراري التالي:

85 - 99	70 - 84	55 - 69	40 - 54	25 - 39	10 - 24	حدود الفئات
30	38	50	42	25	15	التكرار

أ ـ اكتب التكرار النسبي معبرا عنه في نسبة مئوية.
ب ـ اكتب جدول التكرار المتجمع الصاعد، والتكرار النسبي المتجمع الصاعد.
جـ ـ ارسم مضلع التكرار المتجمع الصاعد.

٧) فيما يلي أوزان ستين فأرا (مقاسة إلى أقرب غرام) استخدمت في دراسة تجريبية تتعلق بنقص الفيتامين:

143	117	114	119	123	116	111	106	128	125
125	104	120	132	137	118	121	115	92	136
126	135	133	110	108	129	87	101	115	119
95	120	137	104	82	126	118	110	103	127
113	100	118	126	132	119	105	119	126	146
148	95	113	124	129	117	146	102	125	106

أ ـ لخص هذا البيان الإحصائي في جدول توزيع تكراري متخذا الفئات:

80 - 89; 90 - 99; . . .; 140 - 149

ب ـ ارسم مدرج التكرار، ومدرج التكرار النسبي، ومضلع التكرار.
جـ ـ اكتب جدول التكرار المتجمع الصاعد. والتكرار المتجمع الصاعد النسبي.
د ـ ارسم مضلع التكرار المتجمع الصاعد مستخدما ورقة بيانية.
هـ ـ ما نسبة القياسات التي هي أقل من 109.5؟، أكثر من 89.5؟، أقل من 133؟، وأقل من 105؟
و ـ ما هو القياس الذي يقل عنه ستون بالمئة من القياسات؟ خمس وثلاثون بالمئة من القياسات؟ خمس وعشرون بالمئة من القياسات؟ وخمس وسبعون بالمئة من القياسات؟

٨) مستخدماً فئات طولها 2 مم، اكتب توزيع التكرار وتوزيع التكرار النسبي لسماكة الجلد المعطاة في البيان التالي. (القياسات تمثل سماكة الجلد بالمليمتر في منتصف عضلة الذراع لـ 121 ذكرا).

11.4	15.3	9.1	18.4	10.9	4.7	9.6	20.6	10.4	20.5	22.4	14.3
11.7	11.4	12.7	18.2	15.1	14.6	25.3	11.5	13.2	7.9	12.6	13.9
16.8	11.4	27.3	16.3	13.9	13.2	11.9	20.0	13.2	9.4	18.9	10.7
14.8	17.8	10.8	16.0	15.7	17.7	13.5	11.5	11.1	9.6	15.1	13.6
13.6	8.6	6.9	19.1	18.7	10.1	16.1	20.4	7.9	16.6	18.5	16.2
17.4	18.8	12.6	22.0	9.6	11.1	15.7	23.7	13.3	4.9	8.3	20.1
15.5	23.1	10.2	10.7	15.8	17.6	21.3	16.2	14.9	9.9	9.1	9.9
9.8	8.6	11.8	9.3	14.8	17.3	9.5	13.6	12.4	9.5	14.3	25.7
12.9	22.7	12.1	10.7	16.8	11.3	11.3	11.4	5.9	10.7	14.6	19.8
25.5	7.7	18.4	7.9	7.6	23.3	9.6	8.4	10.4	8.1	12.5	9.1
30.1											

٩) بالعودة إلى المثال (١ ـ ٢)، استخدم الفئات 13.9 - 12.0، 15.9 - 14.0، . . . الخ. لوضع جدول توزيع تكراري لقياسات مستوى الهيموغلوبين في الدم لتسعين عاملا يعيشون في مناطق ترتفع ارتفاعا شاهقا عن سطح البحر.

ارسم مـدرج التكـرار، ومضلع التكـرار، ومضلع التكـرار المتجمع الصـاعـد، واحسب النسبة من القياسات التي تقل عن 16.5 .

١٠) فيما يلي جدول تـوزيع تكـراري لمستوى الهيمـوغلوبين في الـدم لـ 122 عاملا ممن يعيشون في مناطق لا ترتفع كثيرا عن سطح البحر.

حدود الفئات	11.0-11.9	12.0-12.9	13.0-13.9	14.0-14.9	15.0-15.9	16.0-16.9	17.0-17.9
التكرار	6	21	29	43	19	3	1

أ ـ ارسم مدرج التكرار النسبي .
ب ـ ارسم مضلع التكرار النسبي المتجمع الصاعد .
جـ ـ احسب النسبة من القياسات التي تقل عن 16.5 .

١١) فيما يلي جدول توزيـع تكراري للعمر عند الـوفاة مقاسا إلى أقـرب سنة لـ 302 من المرضى الذين توفوا وهم مصابون بالحمى القرمزية :

حدود الفئات	0 -	1 -	2 -	3 -	4 -	5 -	6 -	7 -	8 -	9 -	10 -	15-20
التكرار	18	43	50	60	36	24	22	21	6	5	14	3

ارسم مـدرج التكرار ومضلع التكـرار. ما العمـر الذي تقل عنـه نسبة 90% من حالات الوفاة؟

١٢) فيما يلي جدول توزيع تكراري لحالات الوفاة بسرطان الدم عند الأطفال مصنفة وفقا للعمر مقاسا إلى أقرب سنة . (الولايات المتحدة عام ١٩٧٠م) .

حدود الفئات	(0-0.5	[0.5-1.5)	[1.5-2.5)	[2.5-3.5)	[3.5-4.5)	[4.5-9.5)	[9.5-14.5)
التكرار	68	82	98	137	196	684	434

ارسم مدرج التكرار النسبي .

١٣) فيما يلي قياسات معدل الكولستيرول في الدم لخمسمائة رجل في الأربعينات من عمرهم (49 - 40) مقاسة بالميلليغرام لكل 100 ميلليلتر:

289 385 306 278 251 287 241 224 198 287
275 301 249 288 337 263 260 228 190 282
368 291 249 300 268 283 319 284 205 294
257 256 294 253 221 241 372 339 292 294
327 195 305 253 251 229 250 348 280 378

282 311 193 242 304 270 277 312 264 262
268 251 333 300 250 234 264 291 271 284
322 381 276 205 251 270 254 299 273 252
280 411 195 256 387 241 245 325 289 306
232 293 285 250 260 316 352 309 229 261

272 196 317 188 215 265 266 217 223 354
169 278 188 252 264 314 246 335 377 305
249 318 270 261 324 289 215 228 315 253
262 250 361 304 248 202 284 291 305 261
292 259 369 289 320 287 230 259 321 268

208 386 298 325 262 326 265 281 262 214
277 248 314 279 279 223 202 188 276 261

318	272	245	285	301	234	420	299	255	285
271	283	260	300	308	319	226	235	318	304
291	388	242	277	235	262	176	226	289	247
389	349	210	241	230	260	324	214	296	279
256	260	250	308	294	320	343	312	224	259
305	286	264	209	233	167	272	274	316	291
289	288	175	260	334	248	287	247	222	300
307	269	311	275	273	272	309	307	233	258
263	293	211	263	281	248	349	225	226	388
332	223	186	190	256	321	297	262	380	337
309	227	164	275	283	268	329	259	247	311
246	253	257	328	242	224	283	249	189	207
312	271	277	311	273	316	360	252	243	311
288	226	329	174	248	305	247	309	323	299
174	215	299	183	187	260	268	293	324	325
282	283	324	284	274	285	299	270	354	290
222	280	210	243	199	262	300	218	224	360
293	221	203	386	282	270	277	227	287	226
262	281	319	279	324	279	178	218	246	274
237	239	251	245	337	249	234	202	341	264
281	243	280	346	245	262	213	312	281	312
261	279	356	329	216	326	269	290	300	338
253	284	306	274	277	353	291	333	280	346
270	289	296	296	269	269	275	217	220	351
260	336	323	246	295	296	285	280	330	258
233	219	225	220	210	308	340	319	217	195

262	219	255	278	359	264	273	238	268	301
260	253	237	271	251	226	281	252	338	310

373	217	204	263	246	334	184	222	294	213
331	354	286	291	223	197	324	367	317	253
367	330	315	260	231	266	286	216	286	353
324	315	271	313	306	287	267	274	290	172
275	262	329	283	300	296	238	325	256	244

لاحظ أن كل جزء من الأجزاء العشرة في هذا البيان يتضمن خمسين قياسا .

أ ــ اختر جزءا من الأجزاء العشرة وقم بتصنيفه . ثم ارسم له مدرج تكرار نسبي مستخدما الفئات 189 - 160 ، 219 - 190 ، . . . الخ .

ب ــ ليقم كل اثنين أو ثلاثة من طلاب الفصل بتنفيذ السؤال في جزء محدد من الأجزاء العشرة من البيان وبحيث يتم رسم مدرج تكرار نسبي لكل جزء منها .

جـ ــ قم بضم نتائج الأجزاء العشرة بعضها إلى بعض وارسم مدرج تكرار نسبي للبيان بكامله ، ثم انظر نظرة مقارنة بين مدرجات التكرار النسبي للأجزاء ومدرج التكرار النسبي للبيان بكامله .

١٤) فيما يلي معدل الولادة الخام ومعدل الوفاة الخام في انكلترا وويلز بين 1926 إلى 1976 . وكذلك الفرق بين المعدلين ، ويسمى معدل الزيادة الطبيعية . اكتب جدول التوزيع التكراري لكل منها ، وارسم المضلع التكراري . انظر نظرة مقارنة بين المضلعات الثلاثة . (يمكنك أخذ تسع فئات طول كل منها 1 في معدلات الولادة ومعدلات الزيادة ، وخمس فئات طول كل منها 0.7 في معدلات الوفاة) .

معدل الولادة			معدل الوفاة			معدل الزيادة		
17.8	15.8	14.8	11.6	12.3	12.1	6.2	3.5	2.7
16.7	15.3	14.9	12.3	12.0	12.4	4.4	3.3	2.5
16.7	14.4	15.1	11.7	12.3	11.6	5.0	2.1	3.5
16.3	14.8	14.6	13.4	11.8	12.1	2.9	3.0	2.7
16.3	14.7	14.1	11.4	11.7	14.4	4.9	3.0	0.3
13.9	19.2	15.5	13.5	12.0	12.5	0.4	7.2	3.0
15.6	20.5	15.3	12.3	12.3	11.3	3.3	8.2	4.0
16.2	17.8	15.5	13.0	11.0	11.4	3.2	6.8	4.1
17.2	16.7	15.2	12.7	11.8	11.3	5.0	4.9	3.9
15.9	15.8	15.0	12.6	11.6	11.7	3.3	4.2	3.3
15.7	17.6	17.8	11.7	11.9	11.7	4.0	5.7	6.1
16.1	18.0	17.3	11.5	11.9	11.2	4.6	6.1	6.1
16.4	18.2	16.9	11.7	12.2	11.9	4.7	6.0	5.0
16.5	18.6	16.4	11.6	11.3	11.9	4.9	7.3	4.5
17.1	18.1	16.1	11.5	11.5	11.7	5.6	6.6	4.4
16.0	11.9		11.6	12.0		4.4	- 0.1	
14.8			12.0			2.8		
13.7			11.8			1.9		
13.0			11.8			1.2		
12.2			11.7			0.5		

١٥) فيما يلي أوزان 18645 طفلا مولـودا في جنوب غرب انكلترا (أحياء أو أمـوات) عام ١٩٦٥م مستخدما فئات طولها 1 باوند، اكتب التوزيـع التكراري وتوزيع التكرار النسبي. ارسم مدرجا تكراريا ومضلعا تكراريا لتوضيح البيان.

باوند \ أونزة	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
1	6	1	1	1	3	0	2	2	3	1	3	4	8	2	2	1
2	18	4	2	2	6	2	4	2	10	4	4	2	8	7	4	3
3	14	6	8	5	9	6	8	9	14	2	6	6	7	5	14	7
4	22	14	16	19	16	14	15	19	47	17	23	15	39	30	26	32
5	66	37	42	46	60	41	67	59	106	78	98	68	135	92	106	81
6	323	101	183	157	337	160	205	172	504	215	299	222	496	256	315	228
7	914	225	390	286	697	311	417	291	817	289	369	279	626	246	330	236
8	920	195	292	220	508	200	230	166	485	147	198	110	288	122	146	78
9	395	83	118	72	142	53	69	45	145	35	42	22	91	18	25	10
10	88	12	26	9	23	11	6	4	18	8	7	2	16	4	2	4
11	17	1	3	2	3	1	0	2	2	0	4	1	2	0	1	0
12	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

١٦) تم تنفيذ برنامج استئصال للملاريا في إحدى القرى. وفيما يلي جدول توزيع يعطي النسبة المئوية لقياس الهيموغلوبين في عينة من سكان هذه القرية قبل تنفيذ برنامج الاستئصال. وفي البيان الإحصائي قياسات الهيموغلوبين في عينة أخذت بعد تنفيذ برنامج الاستئصال. اكتب توزيعا مماثلا للبيان الإحصائي الخاص بعينة ما بعد تنفيذ البرنامج. استعن بالرسوم التي تجدها مناسبة.

نسبة الهيموغلوبين	30-39	40-49	50-59	60-69	70-79	80-89	90-99	100-109	المجموع
التكرار	2	7	14	10	8	2	2	0	45
التكرار النسبي (مئويا)	4.4	15.6	31.1	22.2	17.8	4.4	4.4	0	99.9

البيان الإحصائي لعينة ما بعد تنفيذ البرنامج

74	61	51	90	76	71	62	48	80	75	95	75	63	63	43
76	73	69	60	77	80	67	64	53	63	65	74	82	93	103
85	79	66	57	75	89	72	68	50	78	84	69	65	55	91
72	61	74	97	81	99	67	35	52	72	67	87	71	59	70

١٧) استدعت الدراسات التفصيلية لأحد الأمراض في إحدى القرى إجراء حصر شامل للسكان. وفيما يلي التوزيع التكراري لعدد الذكور مصنفين وفقا لشرائح العمر في هذه القرية:

العمر	عدد الذكور	النسبة المئوية ٪
0 - 4	154	18.6
5 - 9	135	16.3
10 - 14	107	12.9
15 - 19	72	8.7
20 - 29	112	13.5
30 - 39	97	11.7
40 - 49	67	8.1
50 - 59	47	5.7
60 - 79	39	4.7
المجموع	830	100.2

أ ـ ارسم مدرج التكرار النسبي لهذا التوزيع.
ب ـ اكتب جدول التكرار النسبي المتجمع الصاعد. وارسم مضلعه.
جـ ـ من الرسم البياني حدد العمر الذي يقسم المجتمع بنسبة 50 - 50. أي ما العمر الذي يمكن القول أن 50% من المجتمع أعمر منه؟

*١٨) فيما يلي عدد الأطباء العاملين وعدد الأسرة في كل من واحد وعشرين من المستشفيات في منطقة الرياض:

* مأخوذ عن التقرير الصحي السنوي الصادر عن وزارة الصحة لعام ١٤٠٦هـ، صفحة ٧٤.

عدد الأطباء	59	85	44	12	12	18	85	51	34	16	28	50
عدد الأسرة	200	266	263	200	160	124	230	205	187	30	72	222

عدد الأطباء	32	34	25	35	43	33	15	30	24
عدد الأسرة	130	115	45	124	146	110	15	100	31

أ ـ اكتب جدول توزيع تكراري لعدد الأطباء متخذا الفئات ., . . . ,41 - 27 ,26 - 12
وجدول توزيع تكراري لعدد الأسرة متخذا الفئات ., . . . ,114 - 65 ,64 - 15 .
ب ـ ارسم مدرج التكرار لكل منهما .

(١ ـ ٦) استخدام بعض الرموز الإحصائية

إن استخدام الرموز للتعبير عن بيان إحصائي ، ومعرفة القواعد التي تخضع لها هذه الرموز، يساعد على التعبير باختصار عن خصائص مهمة للبيان الإحصائي ، واستنباط خطوات العمل الحسابي للوصول إلى القيم العددية لهذه الخصائص . والرمز الأكثر استخداما في الإحصاء هو رمز المجموع Σ [انظر البند (٥) من الملحق]. والقياسات في بيان إحصائي هي ، بصورة عامة ، قيم عددية لمتغير نعبر عنه بحرف x أو y أو z أو أي حرف آخر، وهو يقيس الصفة أو الخاصية التي يدور حولها البيان الإحصائي ، كأن نقيس ، مثلا ، وزنا أو طولا ، أو نسجل عمرا أو معدلا ، أو عدد مرات وقوع شيء معين خلال فترة معينة إلخ ، ويمكننا التعبير رمزيا عن بيان إحصائي لم نحصل عليه بعد ، وإنما نخطط للحصول عليه ، بحروف $x_1 , x_2 , ..., x_n$ حيث n عدد القياسات التي نريد الحصول عليها ، و x_1 هو رمز لأول قياس سنحصل عليه ، و x_2 رمز للقياس الثاني ، وهكذا . . . ، بينما x_n هو رمز لآخر قياس سنسجله . ولو حصل

أن كان العدد الأول الـذي نسجله (عند تنفيذ التجربـة أو جمع البيان الإحصائي) 181، مثلا، فعندئذ نقول إن $x_1 = 181$، وهكذا . . . ، ومن الطبيعي أن يتكـرر حصولنا على القيمة نفسها أكثر من مرة. فلو فرضنا، مثلا، أن $x_1 = x_3 = x_6 = 181$، لقلنا إن القيمة 181 مكـررة ثلاث مـرات. وتجنبا لـلالتبـاس يمكن أن نستخدم حـرفا آخـر y، مثلا، للدلالة على القيم المختلفة التي ورد ذكـرها في البيان، ونكتب في هذه الحالة $y_1 = 181$، ونقول إن y_1 مكررة ثلاث مرات.

وكما نعلم فإن التـوزيع التكراري هـو أي تـرتيب للبيان الإحصـائي يظهر تـوزع قيـاساتـه على فترات معرفـة بصورة اختيـارية تسمى «فئـات». وإذا كان كل عـدد من الأعداد المختلفـة في بيان إحصـائي يمثل فئـة بحد ذاتـه، فسنقول إننـا في حالـة «بيان مرتب» وفيما عـدا ذلك سنقول للتمييـز إننا في حالـة «بيان مصنف» أو «بيـان مبوب». وإذا قمنا بترتيب جملة من القياسات فستأخذ بعد الترتيب الشكل التالي:

جدول (١ ـ ١٣) بيان مرتب

y_i (القيم المختلفة)	y_1	y_2	.....	y_m
f_i (التكرار)	f_1	f_2	.....	f_m

أي أن هناك m قيمة مختلفـة فقط في البيان الإحصـائي الذي يتضمن n قيمة $(n > m)$.
ومن الواضح أنه يمكن التعبير عن مجموع قيم البيان الإحصائي بشكلين متكافئين:

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{m} f_j \; y_j$$

والطرف الأيمن تعبير عن عمليات جمع مكرر للعدد نفسه. فإذا كان القياس 181 مكررا ثلاث مرات، فسيكـون من الأيسر، عند حساب مجموع القياسـات، كتابة 181×3 بدلا من $181 + 181 + 181$. وبصورة عامة، إذا كان أحد القياسات في الطرف الأيسر مكررا f_j

مرة، فقد رمزنـا لهذا القياس المكرر بـ y_j وبدلا من جمع y_j عددا مـن المرات يساوي f_j، كتبنا في الطرف الأيمن $f_j\ y_j$.

أما البيان المصنف فسيأخذ، لأغراض حسابية، الشكل التالي :

جدول (١ ـ ١٤). بيان مصنف

y_i (مركز الفئة)	y_1	y_2	.....	y_m
f_i (التكرار)	f_1	f_2	.....	f_m

وهذا يشير إلى أننـا صنفنا (أو بوبنا) قيم البيـان الإحصائي في m فئة، واعتبرنا مركز كل فئة ممثلا لجميع القياسات التي تنتمي إلى هـذه الفئة، وبذلك استعضنا عن الـ f_j قياسا في الفئة الأولى بمركـز الفئة y_1 واعتبرناه مكررا f_1 مرة واستعضنـا عن الـ f_2 قياسا في الفئة الثانية بمركز هذه الفئة y_2 واعتبرناه مكررا f_2 مرة . . . وهكذا. وعلى سبيـل المثال، لو عدنـا إلى الجدول (١ ـ ٦)، وهـو جدول التوزيع التكـراري لقياسـات حاصل الـذكاء لخمسين مستجدا، وأخـذنا القياسات الفعليـة السبعة ضمن الفئة الـرابعة 101 - 97، لوجدنا أنها :

97, 99, 101, 101, 100, 97, 101

ومجموعها الفعلي هو 696. ولكن الجدول (١ ـ ٦) تلخيص للبيان الإحصائي يغنينا عن العودة إلى مفرداته، حتـى في الحسابات العددية . وإذا أردنا حسـاب مجموع القياسات ضمن هذه الفئة فإننا نتخذ مـركز الفئة، وهو هنا 99، ممثلا لجميع القياسات السبعة، أي نفترض القياسات السبعة في هذه الفئة كأنها :

99, 99, 99, 99, 99, 99, 99

ونعتبر مجموع الفئة مساويا لـ 693 = 99 × 7. ونلاحظ أننا ارتكبنا خطأ بالنقصان قدره 3، وهو الثمن الذي ندفعه في مقابل كفاءة العرض وسهولة وسرعة الحسابات. ومن حسن الحظ أن الأخطاء في الفئات المختلفة لا تكون، عادة، في اتجاه واحد، فلا تكون جميعا أخطاء بالنقصان أو أخطاء بالزيادة، بل تكون في بعض الفئات أخطاء بالنقصان، وفي بعضها الآخر أخطاء بالزيادة، وبذلك يعدل بعضها بعضا، ويكون الخطأ الإجمالي تافها بالمقارنة مع الوفر الكبير الذي حققناه في عملية تصنيف أو تلخيص البيان في هيئة توزيع تكراري، ناهيك عن وضوح العرض وكفاءته سواء في جدول التوزيع التكراري نفسه، أم فيما ينبثق عنه من جداول ورسوم بيانية.

وينبغي أن يكون هذا كافيا لإيضاح نقطة، وهي أن جدولا كالجدول (١ ـ ١١)، نعتمده لحساب خصائص معينة لبيان إحصائي لا يعطي قيم هذه الخصائص بالضبط، كما لو كنا استخدمنا في الحسابات مفردات البيان نفسه، وإنما يعطي تلك القيم بصورة تقريبية، وبفارق زهيد يمكن إغفاله. وفي الجدول (١ ـ ١١) لو رمزنا بـ $\sum_{j=1}^{n} x_j$ لمجموع قياسات البيان الإحصائي الأصلي (قبل التصنيف)، فسيكون المجموع $\sum_{i=1}^{m} f_i y_i$، كما نأخذه من الجدول، قيمة تقريبية للقيمة المضبوطة $\sum_{j=1}^{n} x_j$.

تمرين

احسب مجموع القياسات الخمسين في الجدول (١ ـ ٤) وقارنه مع المجموع الناتج عن استخدام التوزيع التكراري في الجدول (١ ـ ٦).

(١ ـ ٧) مقاييس النزعة المركزية

لا شك في أن الطرق البيانية مفيدة للغاية عند تقديم المعلومات الإحصائية، وأنها تنقل وصفا عاما وسريعا لتلك المعلومات، مما يتفق مع المثل القائل بأن صورة

واحـدة تساوي ألف كلمـة. إلا أن هنـاك حدودا، على أي حـال، لاستخدام الطـرق البيانيـة في مجال وصف وتحليل المعلـومات. وعلى سبيل المثـال، لنفرض أننا نـرغب في مناقشـة البيان الإحصائي أمام مجمـوعة من الناس، وأنـه ليس لدينا طريقـة أخرى غير الطريقـة الشفهية، مما يجعل عـرض المضلع التكراري غير ممكن، ويضطرنـا لاستخدام مقـاييس وصفيـة أخـرى يمكنهـا أن تنقـل إلى المستمعين صـورة ذهنيـة عن المضلع التكراري. والأمر الثاني الذي يضع حدا لاستخدام الطرق البيانية هو صعوبة الاستفادة منهـا في مجال الاستقراء الإحصـائي. وربما اقتصرت فوائدهـا الاستقـرائية على أن يقـدم المضلع التكـراري لعينـة من القيـاسـات نقـوم بتلخيصهـا، تصـورا عن شكل المضلع التكراري للمجتمع من القياسات الذي جاءت منه العينة.

وإذا كنا أمام جملة من القياسات فإن أول مـا تجدر معرفته هو القيمة التي تتمركز عندها القياسات. ومن الملاحظ، مثلا، أنـه في كثير من الظواهر السلوكية والاجتماعية تنزع معظم القيـاسات إلى التمـركز حـول قيمة وسطيـة، فأولئك الذين يتصفـون بحدة شديـدة في المزاج هم قلة وفي المقابل نجـد ذوي المزاج المفرط في برودته قلـة أيضا وذلك قيـاسا على الجمهـرة من النـاس التي تقع بين بين. وأولئك الذين يتصفـون بـالنحافـة الشديـدة يقابلهم أولئك المتصفون بسمانـة مفرطة هم قليلون بـالقياس إلى عامـة الناس التي تحتل مواقعها بين بين. والملاحظة نفسها نجدها سائدة في مجال توزيع الأطوال بين عمالقة وأقزام. فمعظم النـاس في مجتمع بشري معين تميل أطوالها إلى اتخاذ موقع وسط، وقس على ذلك. ولو طبقنا اختبـارا لقياس حاصل الذكـاء على طلاب الجامعة بأسرهم لوجـدنا أن المتفوقين الموهـوبين قلة والمبتلين بالبـلادة قلة، وينـزع حاصل الذكـاء عند معظم الطلبة إلى التمركز حول الوسط.

وفي حياتنـا اليومية، كثيرا ما نستخـدم كلمة «في المتوسط» فنتحـدث عن الرجل «متوسـط الدخل»، والشـاب «متوسط الثقـافة». وقـد يقول أحدنـا: «نادرا مـا أصل متأخرا إلى مقر عملي ونادرا مـا أصل إليه مبكرا، وفي المتوسط يتفق موعـد وصولي تقريبا مع بداية الدوام الرسمي». كما نقول: «إن استهلاكي اليومي من القهوة (أو الشاي) هو في المتوسط كـذا» إلخ. وهذه الاستخـدامات الشائعـة لكلمة متـوسط تعبر عن شعور

داخلي معين يحسه ويفهمه كل منا ولا يستطيع ترجمته بدقة. ومقاييس النزعة المركزية هي محاولة لترجمة هذا الشعور بطريقة دقيقة ومحددة تماما.

وفي لغة الإحصاء يعبر مقياس النزعة المركزية عن القيمة (أو الموضع أو النقطة) التي يتمركز عندها التوزيع التكراري لجملة من القياسات. وعادة ما تحتشد بقية القياسات أكثر ما تحتشد حول ذلك الموضع. وإذ نهتم عادة بمقياس نزعة مركزية لمجتمع من القياسات نلجأ في الغالب إلى عيّنة من المجتمع ونحسب قيمة ذلك المقياس من أجل قياسات العينة ثم نعتبر هذه القيمة التي حصلنا عليها تقديرا أو تخمينا لقيمة المقياس التي نجهلها والخاصة بالمجتمع الذي جاءت منه العينة. وسنستعرض هنا ثلاثة أشكال لقياس النزعة المركزية لجملة من القياسات هي المتوسط والوسيط والمنوال.

(١ ـ ٧ ـ ١) المتوسط (الوسط الحسابي)

والمقياس الأكثر فائدة والأكثر استخداما للنزعة المركزية لجملة من القياسات هو معدلها الحسابي. ويشار إليه غالبا بالوسط الحسابي أو المتوسط.

تعريف المتوسط

متوسط n من القياسات $x_1, x_2, \ldots, x_n$ هو مجموع هذه القياسات مقسوما على عددها. وبصورة رمزية نكتب:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

وإذا لم يكن هناك خشية التباس يمكن كتابة هذه العلاقة على الشكل:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{1}{n}\Sigma x$$

حيث Σx يعني مجموع القيم التي يأخذها المتغير x كافة وعددها n.

مثال (١ ـ ٥)

احسب متوسط القيم 1, 12,14, 6, 3

الحل

$$\bar{x} = \frac{1 + 12 + 14 + 6 + 3}{5} = 7.2$$

ونلاحظ من التعريف مباشرة أن :

$$\sum_{i=1}^{n} x_i = n\bar{x}$$

وفي حالة بيان مرتّب نعبر عن مجموع القياسات على الشكل : (انظر الجدول ١ ـ ١١).

$$f_1 y_1 + f_2 y_2 + \ldots + f_m y_m = \sum_{i=1}^{m} f_i y_i$$

وتصبح العلاقة المذكورة في التعريف السابق للمتوسط كما يلي :

$$\bar{y} = \frac{\sum_{i=1}^{m} f_i y_i}{\sum_{i=1}^{m} f_i}$$

أما في البيانات المصنفة (أو المبوبة) فنفترض أن جميع القياسات التي تنتمي إلى فئة مساوية لمركز هذه الفئة. والخطوات الحسابية ليست إلا تطبيقا للعلاقة الأخيرة من أجل بيان مرتب حيث y_i الآن هي مركز الفئة i، و f_i التكرار الموافق لهذه الفئة. وللتوضيح نأخذ المثال التالي :

مثال (١ ـ ٦)

احسب متوسط حاصل الذكاء في المثال (١ ـ ٣) مستخدما جدول التوزيع التكراري (١ ـ ٦).

الحل

لحساب المتوسط ننظم الجدول التالي :

جدول (١ ـ ١٥). حساب متوسط البيان المصنف في الجدول (١ ـ ٦)

مركز الفئة y_i	التكرار f_i	$f_i y_i$
84	1	84
89	2	178
94	4	376
99	7	693
104	9	936
109	10	1090
114	7	798
119	6	714
124	4	496
المجموع	50	5365

$$\bar{y} = \frac{\sum_{i=1}^{9} f_i \, y_i}{\sum_{i=1}^{9} f_i} = \frac{5365}{50} = 107.30$$

لاحظ أنك عندما تحسب المتوسط من البيان الإحصائي الأصلي في الجدول (١ ـ ٤) ستجد :

$$\bar{x} = \frac{\sum_{i=1}^{50} x_i}{50} = \frac{5364}{50} = 107.28$$

والفرق بين النتيجتين لا يذكر في مقابل الوفر في الجهود الحسابية اللازمة .

(١ ـ ٧ ـ ٢) خواص المتوسط

١ ـ مجموع انحرافات جملة من القياسات عن متوسطها يساوي الصفر.

ولبيان ذلك لنرمز بـ d_i للإنحراف $x_i - \bar{x}$ أي انحراف القياس x_i عن المتوسط $\bar{x}$. ولنحسب مجموع الإنحرافات Σd_i فنجد :

$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n} (x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = n\bar{x} - n\bar{x} = 0$$

وهذه الخاصة توضح الدور المركزي الذي يلعبه المتوسط .

٢ ـ* يكون مجموع مربعات انحرافات القياسات عن قيمة a أصغر ما يمكن عندما يكون $a = \bar{x}$.

لنأخذ مجموع مربعات انحرافات القياسات عن قيمة ما a ، أي $\sum_{i=1}^{n} (x_i - a)^2$ ، فيمكن كتابة :

$$\sum_{i=1}^{n} (x_i - a)^2 = \sum_{i=1}^{n} (x_i - \bar{x} + \bar{x} - a)^2$$

$$= \sum_{i=1}^{n} \left[(x_i - \bar{x}) + (\bar{x} - a)^2\right]$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 + \sum_{i=1}^{n} 2(x_i - \bar{x})(\bar{x} - a) + \sum_{i=1}^{n} (\bar{x} - a)^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 + 2(\bar{x} - a) \sum_{i=1}^{n} (x_i - \bar{x}) + n(\bar{x} - a)^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 + n(\bar{x} - a)^2 \geq \sum_{i=1}^{n} (x_i - \bar{x})^2$$

* البرهان للقراءة فقط .

ذلك لأن $n(\bar{x}-a)^2$ كمية غير سالبة . أي أن مجموع مربعات الانحرافات عن قيمة ما (a) هو دائما أكبر من مجموع مربعات الانحرافات عن المتوسط $\bar{x}$ أو يساويه .

مثال (١ ـ ٧)

في المثال (١ ـ ٥) احسب مجموع الانحرافات عن المتوسط و $\sum_{i=1}^{5}(x_i-7)^2$ و $\sum_{1}^{5}(x_i-7.3)^2$ ثم تحقق من الخاصتين ١ و٢ .

الحل

ننظم جدولا كما يلي :

x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	(x_i-7)	$(x_i-7)^2$	$x_i-7.3$	$(x_i-7.3)^2$
1	− 6.2	38.44	− 6	36	− 6.3	39.69
12	4.8	23.04	5	25	4.7	22.09
14	6.8	46.24	7	49	6.7	44.89
6	− 1.2	1.44	− 1	1	− 1.3	1.69
3	− 4.2	17.64	− 4	16	− 4.3	18.49
المجموع	0	126.80	1	127	- 0.5	126.85

ونلاحظ أن مجموع العمود الثاني صفر بما يتفق مع الخاصية ١ ، وأن كلا من مجموعي العمودين الخامس والسابع أكبر من مجموع العمود الثالث بما يتفق مع الخاصية ٢ .

٣ ـ لنأخذ العلاقة :

$$\bar{y} = \frac{\sum_{i=1}^{m} f_i \, y_i}{\sum_{i=1}^{m} f_i}$$

ولنكتب، للاختصار، n بدلا من $\sum_{i=1}^{m} f_i$. فيمكن إعادة كتابة هذه العلاقة كما يلي:

$$\begin{aligned}\bar{y} &= \frac{1}{n} \sum_{i=1}^{m} f_i \, y_i = \frac{f_1}{n} y_1 + \frac{f_2}{n} y_2 + \ldots + \frac{f_m}{n} y_m \\ &= \omega_1 y_1 + \omega_2 y_2 + \ldots + \omega_m y_m \\ &= \sum_{i=1}^{m} \omega_i \, y_i\end{aligned}$$

حيث $\omega_i = \frac{f_i}{n}$. ويسمى ω_i الوزن الموافق للقياس الـ i، ومجموع هذه الأوزان يساوي الواحد تماما لأن:

$$\sum_{i=1}^{m} \omega_i = \sum_{i=1}^{m} \frac{f_i}{n} = \frac{1}{n} \sum_{i=1}^{m} f_i = \frac{1}{n} \times n = 1$$

ومن الواضح أن كل قياس قد أعطي وزنا يتناسب مع تكرار ظهوره في البيان الإحصائي. ويسمى مثل هذا المتوسط «المتوسط المرجح». ومنه التعريف التالي:

تعريف المتوسط المرجح

لتكن $x_1, x_2, \ldots, x_n$ جملة من القياسات. ولتكن $\omega_1, \omega_2, \ldots, \omega_n$ أعدادا موجبة مجموعها الواحد تماما. فعندئذ يسمى المقدار

$$\bar{x} = \sum_{i=1}^{n} \omega_i \, x_i$$

المتوسط المرجح لهذه الجملة من القياسات. ويسمى ω_i الوزن الموافق للقياس x_i .

مثال (١ ـ ٨)

لنفرض أن درجات طالب في الشهادة الثانوية (الفرع العلمي) منسوبة إلى 100 كانت كما يلي: التربية الإسلامية 87، واللغة العربية 94، واللغة الإنكليزية 97،

والرياضيات 94، والفيزياء 92، والكيمياء 97، والأحياء 98. وأن لكل من التربية الإسلامية والرياضيات ثلاثة أمثال، أما اللغة العربية فلها مثلان، ولكل من المواد الباقية مثل واحد. فاحسب المعدل العام لهذا الطالب؟

الحل

$$\bar{x} = \frac{3\times 87 + 2\times 94 + 1\times 97 + 3\times 94 + 1\times 92 + 1\times 97 + 1\times 98}{3+2+1+3+1+1+1}$$

$$= \frac{3}{12}\times 87 + \frac{2}{12}\times 94 + \frac{1}{12}\times 97 + \frac{3}{12}\times 94 + \frac{1}{12}\times 92 + \frac{1}{12}\times 97 + \frac{1}{12}\times 98$$

$$= 92.92$$

والأوزان $\omega_1, \omega_2, \ldots, \omega_n$ هي، على الترتيب، $\frac{3}{12}$ ، $\frac{2}{12}$ ، $\frac{1}{12}$ ، $\frac{3}{12}$ ، $\frac{1}{12}$ ، $\frac{1}{12}$ ، $\frac{1}{12}$ ومجموعها الواحد.

وتجدر ملاحظة أن تعريف المتوسط هو حالة خاصة من تعريف المتوسط المرجح، حيث الأوزان متساوية، وكل منها يساوي $\frac{1}{n}$ ، ومن الواضح عندئذ أن:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n}\frac{1}{n}x_i = \sum_{i=1}^{n}\omega_i x_i$$

حيث $\omega_i = \frac{1}{n}$. ومجموع الأوزان هو:

$$\sum_{i=1}^{n}\omega_i = \sum_{i=1}^{n}\frac{1}{n} = n\times\frac{1}{n} = 1$$

٤ ـ ليكن $\bar{x}_1$ متوسطا لمجموعة من n_1 قياسا، و $\bar{x}_2$ متوسطا لمجموعة من n_2 قياسا، . . .، و $\bar{x}_m$ متوسطا لمجموعة من n_m قياسا. ولنحسب المتوسط العام لهذه القياسات بعد دمجها في مجموعة واحدة. ولهذه الغاية نطبق تعريف المتوسط فنقول إن المتوسط العام هو مجموع كل القياسات مقسوما على عددها. وإذا لاحظنا أن مجموع المجموعة الأولى هو $n_1\bar{x}_1$ ومجموع المجموعة الثانية هو $n_2\bar{x}_2$ ، . . . ، ومجموع المجموعة الأخيرة هو $n_m\bar{x}_m$ ، يكون:

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + \ldots + n_m \bar{x}_m}{n_1 + n_2 + \ldots + n_m} = \frac{\sum_{i=1}^{m} n_i \bar{x}_i}{\sum_{i=1}^{m} n_i}$$

وهو نـوع من المتـوسط المرجح حيـث $\omega_i = \frac{n_i}{\sum_{i=1}^{m} n_i}$. ونـلاحظ أن الـوزن المعطى لكل متوسط يتناسب مع حجم المجموعة التي يمثلها .

وفي الحالة الخاصة التي يكون فيها $n_1 = n_2 = \ldots = n_m = n$ نجد :

$$\bar{x} = \frac{n \bar{x}_1 + n \bar{x}_2 + \ldots + n \bar{x}_m}{mn}$$

$$= \frac{n(\bar{x}_1 + \bar{x}_2 + \ldots + \bar{x}_m)}{mn}$$

$$= \frac{\bar{x}_1 + \bar{x}_2 + \ldots + \bar{x}_m}{m}$$

وهو متوسط المتوسطات .

واللجوء إلى متوسط المتوسطات عند حسـاب متوسط عام ، هو خطأ شائع ، ولا يصح إلا في حـالـة واحـدة ، هي عنـدمـا يكون كـل منها متـوسطـا للعـدد نفسـه من القياسات .

مثال (١ ـ ٩)

يتألف مقرر الإحصـاء من ثلاث شعب . وقد حسبنـا متوسط عدد أيـام الغياب خلال شهر رجب فكان كما يلي :

الشعبة	الأولى	الثانية	الثالثة
المتوسط	4	5	3

إذا علمت أن أعداد الطلاب في الشعب الثلاث كان 36 في الأولى، و26 في الثانية، و34 في الثالثة، فاحسب متوسط عدد أيام الغياب في مقرر الإحصاء بشعبه الثلاث؟

الحل

مجموع عدد أيام الغياب في الشعبة الأولى = 30 × 4 = 120 يوما،

مجموع عدد أيام الغياب في الشعبة الثانية = 26 × 5 = 130 يوما،

مجموع عدد أيام الغياب في الشعبة الثالثة = 34 × 3 = 102 من الأيام.

المتوسط العام لكافة الشعب = $\frac{120+130+102}{30+26+34} = \frac{352}{90} = 3.91$ يوما.

ونلاحظ أن متوسط المتوسطات $\frac{3+5+4}{3} = 4$ أيام وهو جواب غير صحيح.

(١ ـ ٧ ـ ٣) أثر الانسحاب وتغيير سلم القياس في المتوسط

١ ـ لتكن $x_1, x_2, \ldots, x_n$ جملة من القياسات، متوسطها $\bar{x}$. إذا أضفنا لكل قياس فيها عددا ثابتا c فإن المتوسط يصبح $\bar{x} + c$. ولبيان ذلك نرمز لـ $x_i + c$ بـ y_i، فيكون متوسط القياسات y_i حسب التعريف:

$$\bar{y} = \frac{\sum_1^n y_i}{n} = \frac{\sum_1^n (x_i + c)}{n} = \frac{\sum_1^n x_i}{n} + \frac{\sum_1^n c}{n}$$
$$= \bar{x} + \frac{nc}{n} = \bar{x} + c$$

(كتبنا تسهيلا للطباعة $\sum_1^n$ بدلا من $\sum_{i=1}^n$). ومنه $\bar{x} = \bar{y} - c$. وتسمى مثل هذه العملية أي إضافة عدد ثابت c (قد يكون موجبا أو سالبا) إلى كل قياس، عملية انسحاب [انظر البند (٨) من الملحق ١].

٢ ـ لتكن $x_1, x_2, \ldots, x_n$ جملة من القياسات متوسطها $\bar{x}$. إذا ضربنا كل قياس بعدد ثابت c فإن المتوسط يضرب بالعدد نفسه. ولبيان ذلك، نرمز للعدد cx_i y_i فيكون متوسط القياسات y_i حسب التعريف •

$$\bar{y} = \frac{\sum_1^n y_i}{n} = \frac{\sum_1^n c\,x_i}{n} = \frac{c\sum_1^n x_i}{n} = c\,\bar{x}$$

ومنه $\frac{\bar{y}}{c} = \bar{x}$. وتسمى عملية ضرب كل قياس بعدد ثابت، عملية تغيير في سلم القياس [انظر الفقرة (٨) من الملحق ١].

٣ ـ لتكن $x_1, x_2, \ldots, x_n$ جملة من القياسات متوسطها $\bar{x}$. إذا خضعت هذه القياسات لتحويل وفق العلاقة الخطية :

$$y_i = a x_i + b$$

أي خضعت لعملية تغيير في سلم القياس (الضرب بعدد ثابت a)، ولعملية انسحاب (إضافة عدد ثابت b)، [انظر البند (٨) من الملحق ١]، فالعلاقة نفسها تربط بين المتوسط $\bar{x}$ والمتوسط الجديد $\bar{y}$ أي

$$\bar{y} = a\,\bar{x} + b$$

ولبيان ذلك يكفي أن نكتب :

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{\sum_{i=1}^n (a\,x_i + b)}{n}$$

$$= \frac{\sum_{i=1}^n a\,x_i + \sum_{i=1}^n b}{n} = \frac{a\sum_{i=1}^n x_i}{n} + \frac{n\,b}{n}$$

$$= a\,\bar{x} + b$$

وتستخدم عمليتا الإنسحاب والتغيير في سلم القياس لتسهيل الحسابات. ونوضح الفكرة بالمثال التالي :

مثال (١ ـ ١٠)

يبين الجدول التالي عدد العمال والأجر الأسبوعي الذي يتقاضاه العامل في مستشفى بالريال.

الأجر الأسبوعي	400	600	800	1000	1200	1400	1600	1800	2000
عدد العمال	6	6	4	9	5	3	3	2	1

احسب متوسط الأجر الأسبوعي للعامل في هذا المستشفى.

الحل

إرمز للأجر الذي يدفعه المستشفى بـ x_i وقم بالتحويل التالي من x_i إلى y_i:

$$y_i = \frac{x_i - 1000}{200} = \frac{1}{200}\, x_i - 5$$

تحصل على الجدول التالي:

الأجر الأسبوعي x_i	التكرار f_i	$y_i = \frac{x_i - 1000}{200}$	$f_i y_i$
400	6	−3	−18
600	6	−2	−12
800	4	−1	−4
1000	9	0	0
1200	5	1	5
1400	3	2	6
1600	3	3	9
1800	2	4	8
2000	2	5	10
المجموع	40		4

$$\bar{y} = \frac{\sum f_i \; y_i}{\sum f_i} = \frac{4}{40} = \frac{1}{10} = 0.1$$

ولحساب المتوسط المطلوب نطبق العلاقة :

$$\bar{y} = \frac{\bar{x} - 1000}{200}$$

فنجد :

$$\bar{x} = 200 \; \bar{y} + 1000$$
$$= 20 + 1000 = 1020$$

تمارين (١ ـ ٢)

١) احسب المتوسط لكل مما يلي :

أ ـ 5,2,0,-3,-1

ب ـ 0.004, -0.002, 0.003, 0.001

جـ ـ 2, 2, 3, 7, 10, 100 (لاحظ أثر القيمة 100 على المتوسط) .

٢)* فيما يلي عدد المراكز الصحية والمستوصفات والمستشفيات التي أقيمت في المملكة في كل من الأعوام الثلاثة عشر بين ١٣٩١ هـ و ١٤٠٣ هـ : 4, 14, 36, 47, 26, 92, 127, 48, 31, 67, 79, 11, 22 احسب المتوسط للسنة الواحدة .

٣) متوسط 23 قياسا يساوي 14.7 فما هو مجموع هذه القياسات؟

٤) ابتعنا ستة أنواع من الحاجيات اليومية لمستشفى من كل من ثلاثة مخازن : ا ، ب ، جـ . (الحاجة نفسها من كل مخزن) وكانت الأسعار كما يلي :

---

* مأخوذ من كتاب منجزات خطط التنمية الصادر عن وزارة التخطيط في المملكة . ص ٢٥٤ .

السادسة	الخامسة	الرابعة	الثالثة	الثانية	الأولى	الحاجة
38.50	27.75	23.00	65.75	33.75	16.00	المخزن أ
40.25	29.50	27.50	66.75	40.50	15.00	المخزن ب
41.25	31.50	24.50	68.00	34.00	19.25	المخزن جـ

ما المخزن الذي توصي إدارة المستشفى بالتعامل معه؟

٥) ثلاث مجموعات من القياسات لها متوسطات $\bar{x}_1 = 25$ ، $\bar{x}_2 = 20$ ، $\bar{x}_3 = 22$. وهي تتضمن 20 ، 25، و 30 قياسا ، على الترتيب .

ما هو متوسطها بعد ضمها في مجموعة واحدة؟

٦) معدل أجر الساعة وعدد المستخدمين في مستشفى عند كل من خمس مستويات للأجور كانا كما يلي :

5	4	3	2	1	مستوى الأجور
6.5	6	5.5	5	4.5	معدل أجر الساعة
25	20	15	10	5	عدد العمال

ما معدل أجر الساعة للعامل في هذا المستشفى؟

٧) في المثال (١ ـ ٦) اطرح من مركز كل فئة y_i العدد 107 ، أي اكتب عمودا جديدا $z_i = y_i - 107$. احسب المتوسط $\bar{z}$ ثم تحقق أن $\bar{y} = \bar{z} + 107$.

لاحظ أن طرح 107 من مركز كل فئة جعل العمليات الحسابية أسهل، ويسمى العدد الذي نطرحه «المتوسط الافتراضي». اتخذ العدد 97 متوسطا افتراضيا وأعد العمليات نفسها مستخدما 97 بدلا من 107. هل تجد أنه كلما كان المتوسط الافتراضي أقرب إلى المتوسط الفعلي أصبحت الحسابات أسهل؟

٨) إذا أضفنا 1.4 لكل من القياسات في التمرين ١ فما أثر ذلك على المتوسط؟ أحسب المتوسط الجديد

٩) إذا ضربنا كل قياس في التمرين ١(ب) بألف فما أثر ذلك على المتوسط؟ أحسب المتوسط الجديد.

١٠) المعلومات التالية مأخوذة من سجل لغياب العاملين في عدد من المؤسسات الصحية خلال شهر شعبان:

عدد أيام الغياب y_i	0	1	2	3	4	5	6	7
التكرار f_i	5	15	23	22	17	10	6	3

احسب متوسط عدد أيام الغياب للعامل الواحد.

١١) فيما يلي السجل الدراسي لأحد الطلاب المستجدين في نهاية العام الدراسي ١٤٠٣ ـ ١٤٠٤هـ.

المقرر	١٠٢سلم	١٥١ريض	١٢١ريض	١٠١حين	١٠١ريض	١٠٢عرب	١٠١حص	١٠١سلم	١٠١فيز	١٠١كيم
عدد الساعات	2	3	3	4	3	2	3	2	4	4
التقدير	4.5	3.5	4	3	3	3	4.5	3.5	4	2.5

احسب المعدل التراكمي لهذا الطالب .

١٢) فيما يلي جدول التوزيع التكراري لأعمار خمسين عاملا في إحدى المستشفيات إلى أقرب سنة .

حدود الفئات	15-24	25-34	35-44	45-54	55-64
التكرار	8	11	25	4	2

احسب متوسط العمر للعمال الخمسين في هذا المستشفى .

١٣) في كل من التمارين ٣ إلى ١٧ من مجموعة التمارين (١ ـ ١)، احسب المتوسط .

١٤)* تناولت أنشطة فحص الدم لطفيل الملاريا لعام ١٤٠٦هـ ثماني عشرة منطقة في أنحاء المملكة وكان عدد العينات الإيجابية في كل منها كما يلي :

401, 119, 36, 779, 88, 80, 386, 180, 535,
64, 531, 565, 576, 64, 248, 246, 4331, 81

احسب متوسط عدد العينات الإيجابية للمنطقة الواحدة .

١٥)** فيما يلي عدد المراكز الصحية وعدد الأطباء في كل من المناطق الأربع عشرة في المملكة .

عدد المراكز	232	69	55	90	72	101	26	157	214	45	104	119	69	78
عدد الأطباء	613	234	156	193	145	293	44	355	317	95	180	306	81	130

* مأخوذة من التقرير الصحي السنوي الصادر عن وزارة الصحة لعام ١٤٠٦هـ، ص٢٠٣ .

** مأخوذة من التقرير الصحي السنوي الصادر عن وزارة الصحة لعام ١٤٠٦هـ، ص٤٤ .

أ ـ احسب لكل منطقة متوسط عدد الأطباء في المركز الواحد.
ب ـ احسب متوسط عدد المراكز للمنطقة الواحدة.
جـ ـ احسب متوسط عدد الأطباء للمنطقة الواحدة.
د ـ احسب متوسط عدد الأطباء للمركز الواحد على مستوى المملكة.

(١ ـ ٧ ـ ٤) الوسيط

نلاحظ من دراستنا للمتوسط أنه إذا كان أحد قياسات البيان الإحصائي كبيرا جدا، أو صغيرا جدا، بالمقارنة مع بقية القياسات، تأثر المتوسط كثيرا بهذه القيمة القاصية، ومال إليها، مما يفقد المتوسط الموقع المركزي الذي يفترض أن يشغله. وبالإضافة إلى ذلك فقد رأينا في ختام الفقرة (١ ـ ٢)، أن بعض البيانات يمكن أن تكون وصفية أو ترتيبية ولا يوجد أي معنى لكلمة متوسط، كما عرفناها، في مثل هذه البيانات. وسنعرّف الآن مقياسا للنزعة المركزية يمكن حسابه في كل من البيانات العددية والترتيبية، ومع وجود قيمة قاصية في بيان عددي يمكن أن لا يتأثر أبدا، وفي حال وجود أثر فإنه يكون أثرا طفيفا. ويسمى هذا المقياس الوسيط.

فوسيط n من القياسات هو القياس الواقع في الوسط عند ترتيب هذه القياسات. أي القياس الذي رتبته $\frac{n+1}{2}$ إذا كان عدد القياسات n فرديا، ومتوسط القياسين اللذين رتبتاهما $\frac{n}{2}$ و $\frac{n}{2}+1$ إذا كان عدد القياسات زوجيا.

ملاحظة

في بيان ترتيبي يكون الوسيط هو الصفة المقابلة للقياس الذي رتبته $\frac{n+1}{2}$ في حالة n فردي، أما إذا كان n زوجيا وكان للقياسين الذين رتبتاهما $\frac{n}{2}$ و $\frac{n}{2}+1$ الصفة نفسها فهذه الصفة هي الوسيط، وإذا كانا من صفتين مختلفتين، مثلا أحدهما جيد والذي يليه مقبول، قلنا اصطلاحا إن الوسيط هو بين الجيد والمقبول.

مثال (١ ـ ١١)

ما هو وسيط القياسات

8, 4, 10, 16, 9, 2, 7

الحل

نرتب هذه القياسات فنجد :

2, 4, 7, 8, 9, 10, 16

والوسيط هو القياس الذي رتبته $4 = \frac{7+1}{2} = \frac{n+1}{2}$. أي القياس الرابع . ولكن القياس الرابع في القياسات المرتبة أعلاه هو 8، وبالتالي تكون قيمة الوسيط المطلوبة 8 .

ونلاحظ أن 8 يتوسط مجموعة القياسات ، إذ يقع من القياسات غلى اليمين منه بقدر ما يقع منها على اليسار. كما نلاحظ أننا لم نحتج لأي عمليات حسابية ، إذ قمنا بعملية ترتيب تلتها عملية اختيار.

مثال (١ ـ ١٢)

في فصل يتضمن 9 طلاب كانت التقديرات في أحد الاختبارات كما يلي :

جيد ، ضعيف ، مقبول ، جيد ، جيدجدا ، ممتاز، مقبول ، جيد ، جيدجدا

احسب الوسيط .

الحل

نرتب التقديرات فنجد :

ضعيف ، مقبول ، مقبول ، جيد ، جيد ، جيد ، جيدجدا ، جيدجدا ، ممتاز

والتقدير المقابل للقياس الذي رتبته $5 = \frac{9+1}{2}$ ، أي للقياس الخامس هو جيد ، وهكذا يكون الوسيط في هذا البيان «جيد» .

مثال (١ ـ ١٣)

لدينا القياسات 25 ,22 ,25 ,26 ,37 ,16 ,32 ,26 . ما الوسيط؟

الحل

نرتب هذه الأعداد فنجد :

16, 22, 25, 25, 26, 26, 32, 37

وبما أن عـدد القياسـات $n = 8$ زوجي ، نأخـذ متوسـط القيـاسين اللـذين رتبتاهما $\frac{8}{2} = 4$ و $5 = 1 + 4$. أي العدد الرابع والعدد الذي يليه وهو الخامس. ولكـن العـدد الرابـع هو 25 والعدد الخامس 26، فقيمة الوسيط تساوي :

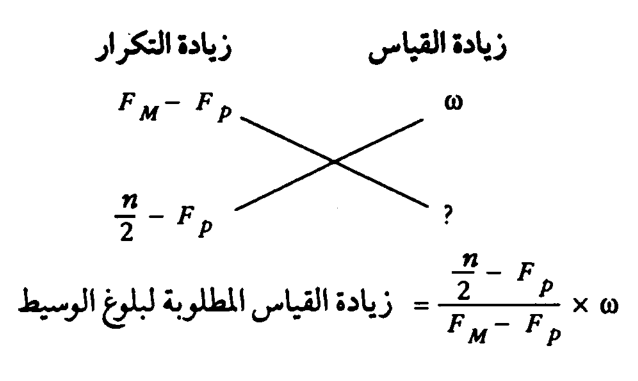
$$\frac{25 + 26}{2} = 25.5$$

ولحساب الوسيط في حـالة بيان مصنف ، ولنرمـز للوسيط بـ M ، نتبع الخطوات التالية بعد كتابة جدول التكرار المتجمع الصاعد :

١ ـ نحسب رتبة الوسيط $\frac{n}{2}$ ، وذلك سواء أكان عدد القياسات n زوجيا أم فرديا .

٢ ـ تحدد رتبة الوسيط الفئة التي ينتمي إليها . ونسميها الفئة الوسيطية ، كما تحدد بالطبع الفئة السابقة للفئة الوسيطية ، وسنسميها اختصارا الفئة السابقة .

٣ ـ لنرمز بـ F_M للتكرار المقابل للفئة الوسيطية في جدول التكرار المتجمع الصاعد ، وبـ F_P للتكرار المقابل للفئـة السابقة . وبـ ω لطول الفئـة ، وبـ L للحد الأعلى الحقيقي المقابل للفئة السابقة . فنجد بعملية تناسب طردي بسيط أن :



ويكون الوسيط إذا :

$$M = L + \frac{\frac{n}{2} - F_p}{F_M - F_p} \times \omega$$

وتجدر ملاحظة أن تصنيف بيان إحصائي يتضمن نوعا من الترتيب لعناصره. ومع أن هذا الترتيب لا يتناول كل قياس بمفرده، إلا أن هناك نوعا من الترتيب الفئوي، إذا صح التعبير. فكل قياس ينتمي إلى فئة هو حتما أصغر من أي قياس ينتمي إلى فئة لاحقة، وأكبر من أي قياس ينتمي إلى فئة سابقة.

(مثال ١ ـ ١٤)

احسب وسيط البيان الإحصائي المصنف المعطى في الجدول (١ ـ ٣).

الحل

نكتب جدول التكرار المتجمع الصاعد فنجد (انظر الجدول ١ ـ ٥).

التكرار المتجمع الصاعد	أقل من	
1	86.5	
3	91.5	
7	96.5	
14	101.5	
23	106.5	الفئة السابقة
33	111.5	الفئة الوسيطية
40	116.5	
46	121.5	
50	126.5	

وتسير الخطوات الحسابية كما يلي:

١ ـ رتبة الوسيط هي $25 = \frac{50}{2} = \frac{n}{2}$ وأول فئة يزيد التكرار المتجمع المقابل لها على 25 تكون الفئة الوسيطية.

٢ ـ نطبق قاعدة التناسب الطردي فنكتب:

زيادة القياس	زيادة التكرار
5	33-23
?	25-23

$$1 = \frac{2}{10} \times 5 = \text{الزيادة المطلوبة لبلوغ الوسيط}$$

$$M = 106.5 + 1 = 107.5 \ (\text{الوسيط})$$

أو نطبق الصيغة التي استخرجناها للوسيط فنجد من الجدول أن :
$F_M = 33$ ، $F_P = 23$ ، $w = 111.5 - 106.5 = 5$ ، $L = 106.5$.

وبالتعويض نجد :

$$M = 106.5 + \frac{25-23}{33-23} \times 5 = 106.5 + 1 = 107.5$$

لاحظ أن الحساب من بيان مصنف هو دائما تقريبي ، ولذلك ترانا تجاوزنا الدقة التامة في حساب رتبة الوسيط فاتخذناها على الدوام $\frac{n}{2}$ سواء أكان n زوجيا أم فرديا . وذلك توخيا للاقتصاد في الجهود الحسابية .

وتجدر ملاحظة أننا إذا رسمنا مضلع التكرار المتجمع الصاعد ومضلع التكرار المتجمع النازل فإن الإحداثي السيني لنقطة تقاطعهما سيكون الوسيط .

تمرين

ارسم على الورقة البيانية نفسها مضلعي التكرار المتجمع الصاعد والنازل للبيان الإحصائي المصنف المعطى في الجدول (١ ـ ٣) واستنتج الوسيط بيانيا .

(١ ـ ٧ ـ ٥) المنوال

رأينا أن المتوسط لا يمكن حسابه إلا من بيانات عددية وأن الوسيط يمكن حسابه من بيانات عددية أو بيانات ترتيبية . وسنعرّف الآن مقياسا للنزعة المركزية يمكن

حسابه في جميع أنواع البيـانات سواء أكانت عددية أم ترتيبيـة أم وصفية . وهذا المقياس يعرف بالمنوال. فالمنوال هو القياس الأكثر تكرارا في جملة من القياسات .

مثال (١ـ١٥)

في تصنيف تناول 2000 من المستجدين في الجامعة حصلنـا على البيان الإحصائي التالي :

	يدخن	لا يدخن
يشرب القهوة	389	1483
لا يشرب القهوة	27	101

ما هو المنوال؟

الحل

المنوال هو «يشرب القهـوة ولا يدخن». فهي الصفة السائدة في هـذه الجملة من القياسات لأن تكرارها 1483 أعلى من تكرار كل من الصفات الثلاث الأخرى .

مثال (١ ـ ١٦)

احسب المنوال في المثال (١ ـ ١٢) .

الحل

المنوال هو تقدير «جيد» باعتباره القياس الأكثر تكرارا .

ملاحظات مهمة

١ ـ المنوال هو الصفة الغالبة في بيان وصفي أو ترتيبي . والصفة الغالبة تعني أنها الصفة التي تتحقق في عـدد من العناصر التي نصنفهـا يفوق عـدد العناصر المحققـة لأية صفة أخـرى . ولا تعني بالضرورة أنها الصفة التي تتحقق في أغلبيـة العناصر أي في أكثر من 50% منها . وقد لا يكون هناك أي صفة تتحقق في أغلبية العناصر.

٢ ـ المنوال هو الصفة أو الصنف الأكثر تكرارا وليس تكرار ذلك الصنف.

٣ ـ التكرار الأعلى لا يعني التكرار الذي يقع بتواتر أكبر ولكن الصفة ذات التكرار الأعلى هي التي تقع بتواتر أكبر.

٤ ـ قد يوجد في بيان وصفي أو ترتيبي صفتان أو وصفان لهما أعلى تكرار (تكرارهما متساويان وكل منهما يمثل التكرار الأعلى بالنسبة إلى بقية الصفات أو الأصناف) فعندئذ يمثل كل منهما منوالا، ونقول إن البيان الإحصائي ثنائي المنوال. والبيان الذي يتضمن منوالا فريدا يسمى وحيد المنوال. وقد يكون هناك ثلاثة أو أربعة منوالات الخ. إلا أنه إذا كان لكل صفة أو صنف التكرار نفسه فنقول عندئذ بعدم وجود منوال ولا نقول إن كل صفة أو صنف هي في حد ذاتها منوال.

وعندما توجد في بيان إحصائي عددي مصنف فئة تتمتع بتكرار أعلى من تكرار أي فئة أخرى و يتناقص التكرار، أو يبقى ثابتا، من فئة إلى أخرى من الفئات السابقة أو اللاحقة لها، نقول إن هذه الفئة هي الفئة المنوالية، ونعتبر مركزها منوالا للبيان الإحصائي.* كما نقول عن التوزيع التكراري لهذا البيان إنه وحيد المنوال أو أحادي المنوال. والمنوال بهذا المعنى هو قمة فريدة في مدرج التكرار موافقة لفئة غير الفئة الأولى وغير الفئة الأخيرة. ومن المستحسن ألا نتحدث عن المنوال باعتباره مقياسا للنزعة المركزية إلا في هذه الحالة. وقد يتضمن المدرج التكراري عدة قمم نسبية. (كل فئة يزيد تكرارها على تكرار الفئة السابقة لها مباشرة، وعلى تكرار الفئة اللاحقة لها مباشرة، تشكل قمة نسبية) وفي حالة وجود قمتين نقول إن التوزيع التكراري ثنائي المنوال. وتكون الفئة الموافقة للقمة الأعلى الفئة المنوالية الرئيسة، ومركزها المنوال الرئيس. وتسمى الفئة الأخرى الفئة المنوالية الثانوية، ومركزها هو المنوال الثانوي. ولا يلعب المنوال، بصورة عامة، دورا كبيرا في علم الإحصاء. ويهتم بالمنوال عادة أصحاب الأعمال

* توجد في بعض الكتب طرق حسابية وصفية لحساب المنوال في مثل هذه الحالة. ولكن حساسيته للتغير في عدد الفئات أو حدودها لا يترك مسوغا قويا لتلك الطرق.

التجارية، والتسـويق والدعاية، والقيمـة الأكثر تكرارا لها مغزى خـاص بالنسبة إليهم فالنـوع الأكثر رواجـا في صناعة معينـة يجذب اهتمام أصحاب هـذه الصناعـة زيادة في إنتاجه ومزيدا من الـدعاية له. كما يهتم به أحيانا الباحثون في العلـوم السلوكية باعتباره قابلا للحساب في جميع أنواع البيانات.

مثال (١ ـ ١٧)

احسب منوال البيان الإحصائي المصنف في الجدول (١ ـ ٦).

الحل

نلاحظ أن أكبر تكرار، وهو 10، يقابل الفئـة [107-111]. وأن التكرار يتناقص عندما نبتعد عن هذه الفئة في كلا الاتجاهين. فهذه الفئة هي إذا الفئة المنوالية، ومركزها 109 هو المنوال.

(١ ـ ٧ ـ ٦) مقارنة بين المتوسط والوسيط والمنوال

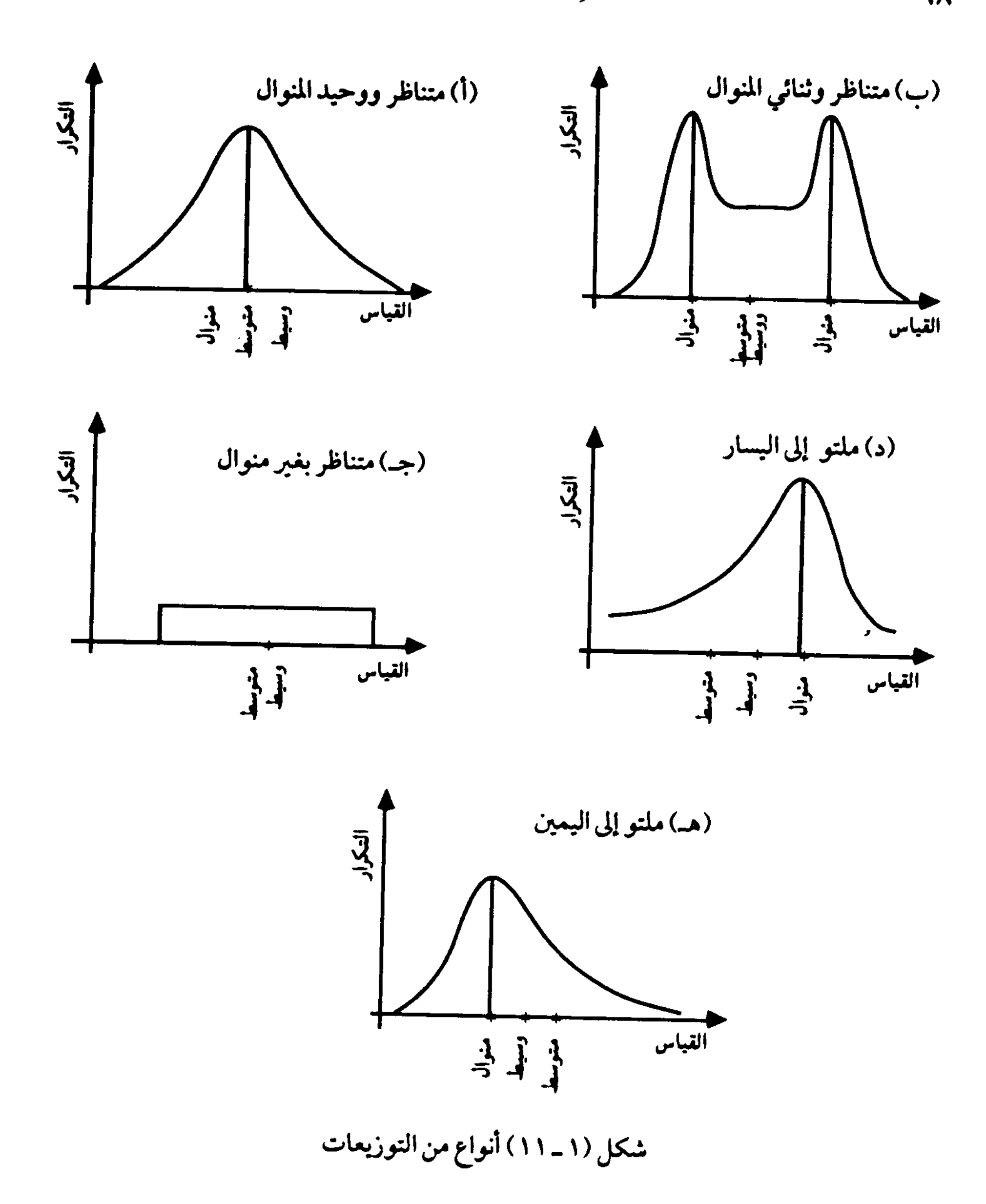
في كل من هـذه المقاييس الثلاثـة محاولة للتعبير عن الموضع الـذي يتمركـز عنده التـوزيع التكـراري، ولذلك سميت مقـاييس النـزعـة المركزيـة. ويتضح من تعـريف المتوسط أن قيمة كل قيـاس من قياسات بيان إحصـائي تسهم في تشكيل قيمة متوسط هذا البيان. ولذلك فقد يتأثر تأثرا بالغـا بالقيم المتطرفة. أما الوسيط فيتحدد من خلال المواقع النسبيـة للقيـاسـات بعضها من بعـض، أي أنه يتحـدد من خـلال رتب هـذه القياسات. ولنأخذ، على سبيل المثال، القياسات 5 ,4 ,3 ,2 ,1 فمتوسطها ووسيطها 3، وإذا أضفنا إليها قيـاسا سادسا كبيرا جدا بـالمقارنة مع بقية القيـاسات، وليكن، مثلا، 69، نجد أن المتوسط أصبح $14 = \frac{84}{6}$، بينما أصبح الوسيط 3.5. فالوسيط زاد بمقدار نصف في حين زاد المتوسط بمقدار 11، والجدير بالذكر أن إضـافة القياس السادس لن تزيد الوسيط إلا بمقدار نصف، مهما كانت قيمته، ولكن زيادة المتوسط ستصبح أكبر فأكبر كلما زادت قيمة القياس السادس الذي أضفناه. أمـا المنوال فلا يتحدد من خلال قيم القياسات، ولا من خـلال رتبها، ولكنه يتحدد من خلال تكرار ظهـورها في البيان الإحصائي.

لنرسم مدرج التكـرار بعناية على ورق مقوى متجانس، ولنـرسم خطا رأسيا من النقطة التي تمثل المتوسط، ثم لنقص الورقة بـدقة على طول محيط المدرج التكراري. ولو أسندنا القطعة الناتجة، وعلى طول الخط الرأسي المرسوم من المتوسط، إلى حرف مستقيم وحاد كحرف سكين لتوازنت. وهذا يعني أن المتوسط هو مركز ثقل التوزيع. ولو رسمنا من القيمة المقابلـة للوسيط خطا رأسيا لقسم المساحة الـواقعة تحت المدرج التكراري إلى نصفين.

وإذا كان المدرج التكراري متناظرا، (متماثلا) تطابقت المقاييس الثلاثة، المتوسط والوسيط والمنـوال. وبهذا المعنى يكون اختلافهـا البيّن كاشفـا عن عدم تناظـر أو التواء حاد في مدرج التكرار أو في مضلع التكرار. وعلى الوجه الآخر، يشير اقترابها من بعضها إلى درجة عالية من التناظر في التوزيع التكراري.

والسؤال الوجيه هنا أي المقاييس الثـلاثة نختاره للتعبير عن الموضع الذي يتمركز عنده التوزيع التكـراري في حال اختلافها عن بعضها؟ والجواب يتـوقف على نوع البيان الإحصائي وعلى شكل التوزيع وعلى الاستخدام الذي نبغيـه للمقياس. ففي حالة بيان وصفي ليس لدينـا إلا المنوال كما ذكرنـا سابقا. وفي بيـانات ترتيبيـة يمكن أن نختار بين المنوال والوسيط أمـا في البيانات العددية فيمكن اختيار أي من المقـاييس الثلاثة. وإذا كان التـوزيع التكراري متناظـرا ووحيد المنوال [انظر الشكـل (١ ـ ١١)أ]، فلا تـوجد مشكلة لأن المقاييس الثلاثـة متطابقة. أما إذا كان التوزيع متنـاظرا وثنائي المنوال كما في الشكـل (١ ـ ١١)ب، فمن الأفضـل أن نقدّم عنـد وصف البيان الإحصـائي كلا من المنوالين، فقد يحجب تقديم القيمة المشتركة للمتـوسط والوسيط نواحٍ مهمة من البيان الإحصائي. فلنفرض مثـلا أننا سألنا 26 من ذوي الـدخل المحـدود عن الحجم الأمثل الذي يتمنـاه لأسرته (عدد الأطفـال مضافـا إليهم الوالدان)، وقـد حصلنا على الجدول التالي:

1	2	3	4	5	6	7	8	9	10	الحجم الأمثل للأسرة
1	2	6	2	1	2	3	6	2	1	التكرار



شكل (١ ـ ١١) أنواع من التوزيعات

ونجد هنا أن المتوسط $\bar{x} = 5.58$ ، وأن الوسيط $M = 6$. وهما تقريبا متساويان وإذا قلنا إن الحجم الأمثل هـو في المتـوسط ، فإننـا نحجـب بـذلك وجـود تيـارين بـارزين بين المستجيبين الستـة والعشرين الذين سألنـاهم ، يمثلهما المنوالان فقيمـة أحد المنـوالين 3 وقيمة المنوال الآخر 8. والتياران الـرئيسان ينقسمان بين من يريد طفـلا واحدا وبين من

يريد عددا من الأطفال يبلغ ستة. وهاتان الناحيتان لا تفصح عنهما القيمة 6 (أي أربعة أطفال). ولا توجد مشكلة في حالة بيان متناظر ليس له منوال كما في الشكل (١ ـ ١١)جـ، فالمنوال غير موجود والمتوسط والوسيط متطابقان.

وفي حالة توزيعات ملتوية نجد أن القياسات في البيان الإحصائي يحتشد بعضها إلى جانب بعض في جانب من المتوسط وتنتشر بعيدا على شكل ذيل في الجانب الآخر منه. ويكون اتجاه الذيل هو اتجاه الالتواء، فإذا كان الذيل على اليسار قلنا إن التوزيع ملتو إلى اليسار كما في الشكل (١ ـ ١١)د. وإذا كان الذيل على اليمين قلنا إن التوزيع ملتو إلى اليمين كما في الشكل (١ ـ ١١)هـ. وفي التوزيعات الملتوية يقع الوسيط دائما بين المنوال والمتوسط. وبما أن المنوال بالطبع عند القمة فالمتوسط يأخذ موقعه في الجانب الآخر أقرب إلى الذيل. وهذا يرشح الوسيط مقياسا أكثر استقرارا وأفضل تعبيرا عن الموقع الذي يتمركز عنده التوزيع. فالمتوسط كما رأينا شديد الحساسية للقيم المتطرفة، ولذلك نراه مائلا إلى اتجاه الذيل. أما المنوال فهو دائما في جانب القمة وشديد الحساسية، في البيانات المصنفة، للتغير في عدد الفئات أو حدودها مما يجعله أيضا خارج الإعتبار. وهكذا نفضل الوسيط في البيانات التي تتصف بالتواء واضح. ولتوضيح هذه الميزة للوسيط لنفرض أن مؤسسة تدفع رواتب سنوية لموظفيها ومستخدميها بالريال كما يلي:

180000, 72000, 30000, 18000, 3000, 3000, 3000, 3000

فنجد في هذا البيان أن المتوسط = 39000 ريال، وأن الوسيط = 10500 ريال، وأن المنوال = 3000 ريال. ومن الواضح أن الأرقام الثلاثة بتعبيرها عن متوسط الرواتب السنوية في هذه المؤسسة تقدم انطباعات مختلفة تماما. وأن كلا من المنوال والمتوسط لا يعبران بموضوعية عما يجري في المؤسسة. ولو أن مراقبا من وزارة الشؤون الاجتماعية والعمل أراد أن يظهر المؤسسة بمظهر الذي يدفع رواتب متدنية جدا في المتوسط لاختار المنوال مقياسا للنزعة المركزية. وفي المقابل فإن مدير المؤسسة سيختار المتوسط وهو 39000 ريال ليثبت أن رواتب الشركة مرتفعة. أما الباحث الاجتماعي الذي يرغب في التعبير

بموضـوعية أكثر عما يجري فعـلا في الشركة فسيختار الـوسيط وهو 10500 ريال مقياسا للنزعـة المركزية . وهذا المثـال يوضح أيضا سبب قـولنا إن الاختيار بين المقاييس الثـلاثة يتوقف أحيانا على الغرض الذي نريده من المقياس .

وإلى جانب هذه الميزة للوسيط في التوزيعات الملتوية يمكن أن نضيف أنه بصورة عامة سهل الحساب وغير حساس للقيم المتطرفة ويبقى حسـابه ممكنا في بيانات ناقصة سقطت منهـا قيم بعض المفردات المتطـرفـة التي نعرف مـواقعها النسبيـة . وعلى سبيل المثال ، لنأخذ البيان التالي عن درجات سبعة طلاب :

71, 68, – , 75, – , 77, –

ففي هـذا البيـان ثلاث درجات غير معروفـة.ولكن لنفرض أننا نعلم عن الطلاب الثلاثة الذين لا نعلم بـالتحديد درجـاتهم أن أحدهم راسب ، والآخـر ناجح بتقديـر مقبول ، والثالث ناجح بتقدير ممتاز. فيمكننا معرفة الوسيط ، ويمكن ترتيب معلوماتنا كما يلي :

ممتاز, 77, 75, 71, 68 , مقبول , راسب

واستنتاج أن الوسيط هو 71. لا بل أكثر من ذلك لـو علمنا أن ثـلاثة منهم بين راسب ومقبـول وثـلاثة نـالـوا «جيـد مرتفع» أو أفضل ، وأن أحـدهم نـال 71، لكانـت هذه المعلومات كافية لاستنتاج أن الوسيط 71 .

ويبقى المتوسط مقياسـا للنزعة المركزية يتمتع بخصائص مهمـة تجعله مستخدما على نطاق واسع في علم الإحصاء وستتضح هذه النقطة للقارىء عبر هذا الكتاب ، ولو أخـذنا عينـات مختلفة بـالحجم نفسه مـن مجتمع وحسبنا لكل عينـة المتوسط والـوسيط لوجدنا أن التغير من عينة إلى أخـرى هو أقل في قيم المتوسط منه في قيم الوسيط ، ونعبر عن ذلك بقولنـا إن المتوسط أكثر استقـرارا من الوسيط عبر عينات متكـررة نسحبها من مجتمع معيّن .

تمارين (١ ـ ٣)

١) أوجد الوسيط لكل من المجموعات التالية من القياسات :

أ - 6, 4, -1, 5, 1, 2

ب - 17.2, 16.9, 17.5, 16.4, 17.1

جـ - 2, 2, 1, 1, 1, 3, 1, 2, 1, 1

٢) في التمرين ٩ من مجموعة التمارين (١ ـ ٢)، احسب الوسيط والمنوال لعدد أيام الغياب.

٣) في التمرين ١١ من مجموعة التمارين (١ ـ ٢)، احسب الوسيط والمنوال لأعمار العمال الخمسين في المستشفى. أي المقاييس الثلاثة تفضل؟

٤) صنفنا عينة من محصول التفاح وفقا لوزن التفاحة مقاسا بالأونزة، فحصلنا على التوزيع التكراري التالي :

الوزن	0.21-0.30	0.31-0.40	0.41-0.50	0.51-0.60	0.61-0.70
التكرار	31	45	36	23	11

احسب المتوسط والوسيط والمنوال. أيهما تفضل للتعبير عن النزعة المركزية؟

٥) احسب المتوسط والوسيط إذا علمت توزيع التكرار النسبي التالي :

الفئة	21 - 40	41 - 60	61 - 80	81 - 100
التكرار النسبي	0.24	0.36	0.28	0.12

٦) يبين الجدول التالي توزع فترة الإقامة في المستشفى لأطفال تحت سن الخامسة عشرة

من العمر ممن أجروا عمليات استئصال اللوز والزوائد الأنفية، وذلك في كل من أربع مجموعات من المستشفيات.

احسب المتوسط والوسيط والمنوال لطول فترة الإقامة في كل مجموعة من المستشفيات.

المجموع	فترة الإقامة (بالأيام)											مجموعة المشافي
	10	9	8	7	6	5	4	3	2	1	0	
178	1	-	1	2	4	5	36	113	16	-	-	A
33	-	-	-	-	27	-	2	2	1	1	-	B
146	6	4	1	7	35	28	20	33	12	-	-	C
184	4	1	2	27	11	28	6	2	6	97	-	D

٧) يبين الجدول على الصفحة التالية جزءا من دراسة لحجم رد الفعل لاختبار الليبرومين في ثلاث جماعات من الأطفال، وقد أعطي الأطفال في الجماعة الأولى لقاح الـ .B.C.G عند ولادتهم عن طريق الفم، وفي الجماعة الثانية أعطي الأطفال اللقاح عن طريق أدمة الجلد، ولم يعط أطفال الجماعة الثالثة لقاح الـ .B.C.G.

احسب لكل جماعة المتوسط، الوسيط والمنوال لقياسات نصف قطر رد الفعل.

٨) حشرات من نوع العت تتغذى على فئران مصابة بدودة الفيلاريا. وقد أخذت هذه الحشرات بعد فترة وأحصي عدد الميكروفيلاريا في كل عت. ويمثل البيان التالي نتائج التعداد لخمسين عتا. احسب المتوسط، الوسيط، والمنوال لهذه القياسات وعلق على الفروق بين هذه المقاييس الموضعية الثلاثة.

7	12	3	3	1	8	0	7	2	0
10	15	3	19	1	2	2	15	3	4
7	0	9	0	18	4	6	6	10	1
1	9	14	3	7	5	7	5	14	20
6	1	2	14	3	3	5	1	4	3

٩) فيما يلي مستوى السكر في الدم مأخوذا في الصباح قبل تناول الفطور لعشرة أطفال :

56, 62, 63, 65, 65, 65, 68, 70, 72

احسب الوسيط والمنوال .

١٠) يبين الجدول التالي التكرار النسبي المتجمع لعمر العروس وفي أربع عينات من النساء مأخوذة من أربع جماعات ، تاريخ الميلاد في الجماعة الأولى يعود إلى ما قبل 1925 ، وفي الثانية بين 1925 إلى 1934 ، وفي الثالثة من 1935 إلى 1944 ، وفي الرابعة من 1945 إلى 1954 .

ارسم مضلع التكرار النسبي المتجمع الصاعد لكل من العينات الأربع ، وقدر العمر الوسيط للعروس في كل جماعة .

نصف قطر رد الفعل (بالميلليمتر)	عدد الأطفال		
	B.C.G. عن طريق الفم	B.C.G. تحت الجلد	لم يعط B.C.G.
1	-	2	7
2	-	3	2
3	36	53	39
4	22	22	-
5	29	42	9
6	18	15	2
7	10	4	-
8	8	4	2
9	3	-	-
10	3	2	-
11	-	-	-
12	3	-	-
13	-	-	-
14	2	1	-
15	1	-	-
16	1	-	-
المجموع	136	150	61

العمر عند الزواج	قبل 1925 N = 61	1925-1934 N = 83	1935-1944 N = 90	1945-1954 N = 106
	%	%	%	%
9-10	3.4	6.0	5.6	9.6
11-12	6.9	13.3	18.0	21.1
13-14	39.7	27.7	38.2	39.4
15-16	58.6	56.2	68.5	73.1
17-18	63.8	74.7	77.5	90.3
19-20	74.1	80.7	85.4	98.8
21-22	79.3	86.7	89.9	99.0
23-24	82.8	88.0	95.5	100
25-26	87.9	90.4	97.7	
27-28	89.7	92.8	97.7	
20-30	93.1	96.4	98.8	
> 30	100	100	100	

١١) فيما يلي أوزان عشرة حيوانات تجربة وذلك بعد مداخلة جراحية (مقاسة بالكغ):

13.2, 15.4, 13.0, 16.6, 16.9, 14.4, 13.6, 15.0, 14.6, 13.1

احسب الوسيط.

١٢) فيما يلي المسافة (إلى أقرب ميل) التي قطعها كل من خمسة عشر مريضا حتى وصلوا إلى أقرب مستوصف:

5, 9, 11, 3, 12, 13, 12, 6, 13, 7, 3, 15, 12, 15, 5

ما وسيط المسافة التي يقطعها المريض حتى يصل إلى أقرب مستوصف؟

١٣) كانت فترة الإقامة بالأيام لأول أحد عشر مريضا أدخلوا إلى جناح للأمراض النفسية افتتح حديثا في أحد المستشفيات كما يلي:

29, 14, 11, 24, 14, 14, 28, 14, 18, 22, 14

احسب الوسيط والمنوال لعدد أيام الإقامة في المستشفى.

١٤) فيما يلي جدول توزيع تكراري يلخص بيانا إحصائيا عن درجة تلوث الهواء (مقاسة بالميكروجرام في المتر المكعب) في 57 مدينة كبيرة.

الفئات	10-19	20-29	30-39	40-49	50-59	60-69	70-79
التكرار	5	19	10	13	4	4	2

احسب المتوسط والوسيط والمنوال.

١٥) في كل من التمارين ٣ إلى ١٧ من مجموعة التمارين (١ ـ ١)، احسب الوسيط.

١٦) بعد افتتاح مركز حديث للتسوق قرب ضاحية معينة، ازدادت حركة المرور فيها، وقد حددت إدارة المرور السرعة القصوى في شارع الضاحية بـ 35 كم/ سا. وبعد شكاوي عن عدم التزام السيارات بهذا الحد قامت دورية مرور خلال 15 دقيقة من المراقبة برصد سرعات 25 سيارة مرت من ذلك الشارع. وحصلت على البيان التالي:

15, 40, 47, 25, 37, 23, 20, 38, 29, 40,
35, 28, 37, 38, 35, 37, 27, 36, 30, 38,
40, 43, 25, 20, 42

أ ـ إذا كنت من سكان الضاحية الذين يرغبون في استخدام هذا البيان لإثبات أن السيارات بصورة عامة لا تتقيد بحد السرعة المفروض، فهل تستخدم المنوال، أم الوسيط أم المتوسط؟

ب ـ إذا كنت ممن يعارضون فرض حد للسرعة وتريد استخدام هذا البيان لدعم وجهة نظرك بأن السيارات ملتزمة بصورة عامة بلوحة المرور، أي المقاييس تختار؟

١٧) تهتم شركة بمعرفة مدى استخدام موظفيها للهواتف في مكالمات شخصية. وفي أحد الأيام راقبت عدد المكالمات الشخصية التي قام بها كل موظف فحصلت على البيان التالي:

عدد المكالمات الشخصية	10	9	8	7	6	5	4	3	2	1	0
عدد الموظفين	1	1	0	0	11	8	6	23	112	65	273

أ ـ أوجد متوسط المكالمات الشخصية للموظف الواحد في ذلك اليوم.

ب ـ آخذا في اعتبارك أولئك الذين استخدموا الهاتف لأغراض شخصية فقط، ما المقياس الذي تجده أفضل تعبيرا عن النزعة المركزية؟ احسب هذا المقياس.

١٨) في التمرين ٥ من مجموعة التمارين (١ ـ ١)، احسب المتوسط والوسيط والمنوال. أي المقاييس الثلاثة تفضل ولماذا؟

١٩) في التمرين ١٨ من مجموعة التمارين (١ ـ ١)، احسب مقياس النزعة المركزية الذي تعتقد أنه مناسب في كل من بيان الأطباء وبيان الأسرّة. وأوضح أسباب تفضيلك.

٢٠) في التمرين ١٣ من مجموعة التمارين (١ ـ ٢) احسب الوسيط. أيهما تفضل المتوسط أم الوسيط ولماذا؟

٢١) فيما يلي بيان بعدد الزيارات التي قام بها المراجعون للعيادات الخارجية بالمستشفيات ومراكز الرعاية الصحية الأولية في المناطق الأربع عشرة في المملكة، وذلك خلال عام ١٤٠٦هـ، بآلاف المراجعين:

11169, 4330, 4870, 3029, 2050, 4802, 1577,
6375, 6034, 1480, 3876, 3465, 2826, 1794.

احسب المتوسط والوسيط أيهما تعتقد أنه الأفضل لقياس النزعة المركزية ولماذا؟

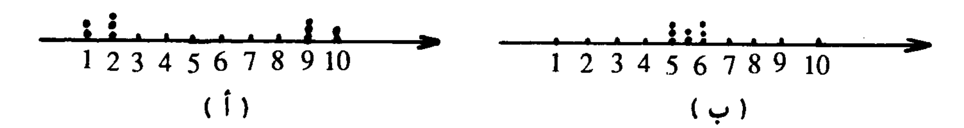
(١ ـ ٨) مقاييس التشتت

ناقشنا في الفقرات السابقة معايير موضعية تهدف إلى تحديد الموضع الذي تتمركز عنده جملة من القياسات الإحصائية. ولكنها لا تكفي وحدها لتشكيل صورة ذهنية متكاملة عن التوزيع التكراري للبيان الإحصائي. وإلى جانب المكان الذي يشكل مركز التوزيع نحتاج إلى معرفة كل ما يمكن معرفته عن خاصية التغير من قياس إلى آخر ضمن البيان الإحصائي، وعن مواقع القياسات بالنسبة إلى مركزها. فالقياسات 1, 2, 3, 4, 5 لها متوسط يساوي 3 وهو بالذات متوسط للقياسات 602 ,209 ,4 ,200 - ,600 - . ولكن شتان ما بين المجموعتين من القياسات من حيث درجة تجمعها حول المركز المشترك لكل منها وهو 3. ومن المعروف أنه لا يمكن لقياسات بيان إحصائي أن تكون متساوية. ولو قسنا، مثلا، أطوال مجموعة من أوراق نبات معين، لاختلف القياس من ورقة إلى أخرى، ولو كان الشخص نفسه هو الذي يكرر قياس ظاهرة معينة مستخدما الجهاز نفسه في كل مرة، فسيختلف القياس الذي يحصل عليه من محاولة إلى أخرى. والتغير ظاهرة ملازمة لكل بيان إحصائي، وإذا كان التوزيع التكراري للبيان الإحصائي يتمركز عند المتوسط الحسابي، فهو ينتشر على جانبي هذا المتوسط، وكلما كان التغير كبيرا من قياس إلى آخر، اتسع انتشار القياسات حول متوسطها. وبصورة عامة. فإن القياسات التي تحتشد وتتجمع حول متوسطها، وقريبا منه، يكون تشتتها صغيرا. بينما يكون تشتت القياسات المبعثرة التي تنتشر بعيدا على جانبي المتوسط، تشتتا كبيرا. وسنحاول فيما يلي تقديم معايير كمية لقياس شدة تبعثر القياسات في بيان إحصائي، أو لقياس درجة انتشار وتشتت القياسات حول متوسطها. وسنبدأ بتعريف المدى.

(١ ـ ٨ ـ ١) تعريف المدى

مدى بيان إحصائي هو الفرق بين أكبر قياس وأصغر قياس في البيان الإحصائي.

ومن الواضح أن المدى يعطي فكرة واضحة عن المسافة على محور الأعداد التي يتوضّع فيها البيان الإحصائي. وإذا استثنينا القيمتين المتطرفتين في البيان الإحصائي فإن المدى بمفرده عاجز عن تقديم أية معلومات عن أسلوب انتشار بقية القياسات حول المتوسط. وعلى سبيل المثال، إذا كان لدينا عشرة قياسات متوسطها 5.5 وأحدها 1 وأكبرها 10. فيمكن تصور هذه القياسات العشرة بأشكال عديدة تختلف اختلافا شديدا في درجة تبعثرها وتشتتها حول المتوسط، ونجد في الشكل (١ ـ ١٢) تصورين مختلفين. ففي الشكل (١ ـ ١٢)أ، نجد القياسات 10 ,10 ,9 ,9 ,9 ,2 ,2 ,2 ,1 ,1 وفي الشكل (١ ـ ١٢)ب، نجد القياسات 10 ,6 ,6 ,6 ,5.5 ,5.5 ,5 ,5 ,5 ,1 ومع أن للمجموعتين المدى نفسه وهو 9 = 1 - 10، إلا أن الفارق كبير بين درجة تمركز كل منهما حول المتوسط المشترك 5.5 :



شكل (١ ـ ١٢)

وإذا كان المدى يضم بين طرفيه جميع قياسات البيان الإحصائي فلماذا لا نفكر بمدى أكثر تواضعا يضم بين طرفيه نسبة عالية من القياسات (ثمانين بالمائة منها مثلا) بدلا من أن يضمها جميعها. ولو عرفنا مثلا، القياس الذي يقل عنه 10% من القياسات، وسنسميه المئين عشرة، والقياس الذي يقل عنه 90% من القياسات، وسنسميه المئين تسعين، فبين المئين عشرة والمئين تسعين يقع ثمانون بالمائة من القياسات. ولو حسبنا هذين القياسين ووجدناهما قريبين من بعضهما، فسيعطينا ذلك تصورا مفيدا تماما عن واقع انتشار أو تشتت البيان الإحصائي. وسنعرف فيما يلي المئينات باعتبارها وسيلة من وسائل التعبير عن تشتت بيان إحصائي.

(١ ـ ٨ ـ ٢) تعريف المئينات

ليكن r أي عـدد صحيح بين الصفر والمائة، نعـرف المئين r لبيان إحصائي بأنه العدد الذي يقل عنه r بالمائة من قياسات البيان الإحصائي .

ونلاحظ من هـذا التعريف أن المئين r ، وسنرمز لـه بـ P_r ، هو القياس المقابل لتكرار متجمع يساوي $n \times \frac{r}{100}$ ، حيث n عدد القياسات . ومن الواضح أن P_{50} هو الوسيط باعتباره القياس المقابل لتكرار متجمع يساوي $n \times \frac{50}{100}$ أو $\frac{n}{2}$. ويسمى P_{25} (المئين 25) الربيع الأدنى، باعتباره يحصر إلى اليسار منه ثلاثة أرباع القياسات وسنرمز له بـ Q_1 . كما يسمى P_{75} (المئين 75) الربيع الأعلى باعتباره يحصر إلى اليسار منه ثـلاثة أرباع القياسات وسنرمز له بـ Q_3 . والمسافة بين هذين القياسين، أي الفرق بين الربيع الأعلى والربيع الأدنى، تسمى المدى الربيعي .

$$\text{المدى الربيعي} = Q_3 - Q_1$$

ويضم المدى الـربيعي بين طرفيـه *50%* من القيـاسـات . ويعتبر نصف المدى الربيعي $\frac{Q_3 - Q_1}{2}$ أحد معايير التشتت .

وتجدر ملاحظـة أن أكبر قياس في البيـان الإحصائي هـو المئون 100 (P_{100}) وأن أصغر قياس هو المئون صفر(P_0) . وأن المدى هو $P_{100} - P_0$.

ولقد أوضحنـا عمليـا طريقـة حساب أي مئين في الفقـرة (١ ـ ٤)، ولحسـاب (P_r) ، بصورة عامة ، نكتب أولا جدول التكرار المتجمع الصاعد، ثم نحسب رتبة المئين r وهـي $n \times \frac{r}{100}$ ، حيـث n عدد القياسات في البيان الإحصائي . وتحدد رتبة المئين الفئة التي ينتمي إليهـا المئون وسنسميها فئـة المئين، كما تحدد بالطبع الفئة السـابقة لها . لنرمز الآن بـ F_P للتكرار المقابل لفئة المئين في جدول التكـرار المتجمع الصاعد، وبـ F_b

للتكرار المقابل للفئـة السابقة، وبـ w لطول الفئة، وبـ L للحد الأعل الحقيقـي المقابل للفئة السابقة. فنجد بعملية تناسب طردي بسيط أن:

زيادة القياس	زيادة التكرار
w	$F_P - F_b$
?	$\frac{nr}{100} - F_b$

ومنه:

$$\text{زيادة القياس المطلوبة لبلوغ المئين } r = \frac{\frac{nr}{100} - F_b}{F_P - F_b} \times w$$

ويكون المئون r المطلوب:

$$P_r = L + \frac{\frac{nr}{100} - F_b}{F_P - F_b} \times w$$

مثال (١ ـ ١٨)

احسب Q_1 (الربيع الأدنى)، و Q_3 (الربيع الأعلى) للتوزيع التكراري في الجدول (١ ـ ٣) واحسب نصف المدى الربيعي.

الحل

١ ـ رتبة الربيع الأدنى هي $n \times \frac{25}{100} = 50 \times \frac{25}{100} = 12.5$

وأول فئة يزيد التكرار المتجمع المقابل لها على 12.5 تكون فئة الربيع الأدنى.

٢ ـ نطبق قاعدة التناسب الطردي فنكتب:

نكتب جدول التكرار المتجمع الصاعد وتسير الخطوات الحسابية كما يلي :

التكرار المتجمع الصاعد	أقل من	
1	86.5	
3	91.5	
7	96.5	— الفئة السابقة
14	101.5	— فئة الربيع الأدنى
23	106.5	
33	111.5	— الفئة السابقة
40	116.5	— فئة الربيع الأعلى
46	121.5	
50	126.5	

زيادة القياس	زيادة التكرار
5	14 - 7
?	12.5 - 7

الزيادة المطلوبة لبلوغ الربيع الأدنى $= \frac{12.5-7}{14-7} \times 5 = 3.93$

$$Q_1 = 96.5 + 3.93 = 100.43$$

ولحساب الربيع الأعلى (Q_3) نجد بصورة مماثلة أن رتبة الربيع الأعلى هي $50 \times \frac{75}{100} = 37.5$

زيادة القياس	زيادة التكرار
5	40 - 33
?	37.5 - 33

الزيادة المطلوبة لبلوغ الربيع الأعلى $= \frac{37.5-33}{40-33} \times 5 = 3.21$

$$Q_3 = 111.5 + 3.21 = 114.71$$

أو نطبق الصيغة العامة التي استخرجناها من أجل المئينات فنجد من الجدول ، في حالة الربيع الأدنى أن $r = 25$, $F_b = 7$, $F_p = 14$, $L = 96.5$, $\omega = 5$ ثم نعوض في العلاقة :

$$P_r = L + \frac{\frac{nr}{100} - F_b}{F_Q - F_b} \times \omega$$

وفي حالة الربيع الأعلى يكون $[\ \omega = 5, L = 96.5, F_p = 14, F_b = 7, r = 25\ ]$.

وأخيرا:

$$\text{نصف المدى الربيعي} = \frac{114.71 - 100.43}{2} = \frac{14.28}{2} = 7.14$$

ولكن لماذا لا نبحث عن مقياس للتشتت يسهم في تشكيله كل قياس من قياسات البيان الإحصائي بدلا من أن يقتصر على مئينين أو على أكبر قياس وأصغر قياس؟ ومن الواضح أن التشتت يعود في الأساس إلى قرب أو بعد القياسات عن متوسطها. فلنحاول إذاً التعبير عن التشتت بدلالة انحراف كل قياس عن المتوسط، أي بدلالة $d_i = x_i - \bar{x}$. ونعلم من خواص المتوسط أن مجموع هذه الانحرافات يساوي الصفر مما لا يترك مجالا للتفكير في متوسط هذه الانحرافات كمقياس للتشتت. ولكن حل هذه المشكلة سهل طالما أنه يعود إلى وجود انحرافات موجبة وانحرافات سالبة، فلماذا لا نحسب متوسط القيم المطلقة للانحرافات؟

(١ ـ ٨ ـ ٣) تعريف متوسط الانحراف

لتكن $x_1, x_2, \ldots, x_n$ جملة من القياسات متوسطها $\bar{x}$. نعرف متوسط الانحراف لهذه القياسات، ونرمز له بـD، بأنه متوسط القيم المطلقة لانحرافات القياسات عن متوسطها.

$$D = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}| = \frac{1}{n}\sum_{i=1}^{n}|d_i|$$

وتلافيا للتعقيدات التي يسببها وجود القيمة المطلقة عند استخدام المعيار D في التحليل الإحصائي، يمكن اللجوء إلى حل آخر لمشكلة الانحرافات الموجبة والانحرافات السالبة، وذلك بأخذ مربعات الانحرافات بدلا من قيمها المطلقة، مما يؤدي إلى تعريف مقياس للتشتت يسمى التباين.

(١ ـ ٨ ـ ٤) تعريف التباين

تباين مجتمع من القيـاسات يتضمن N قياسا . x_1 , x_2 , ... , x_N هو متـوسط مربعات انحرافات القياسات عن متوسطها . وسنرمز له بـσ^2 ، وبصورة رمزية نكتب :

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2$$

وتجدر ملاحظـة أن التباين مـوجب دوما لأنـه ناشيء عن مجمـوع مربعـات ، أي مجموع كميات موجبة . ويكون التباين صفرا إذا وفقط إذا كانت القياسات جميعها متساوية .

(١ ـ ٨ ـ ٥) الانحراف المعياري لمجتمع

الانحراف المعياري σ هو الجذر التربيعي الموجب للتباين

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2}$$

ويقاس الانحراف المعياري بوحدة القياس نفسها المستخدمة في البيان الإحصائي .

ملاحظة

الرمز المستخدم σ هو الحرف الأبجدي سيجما في الأبجدية اليونانية بالخط الصغير ويكتب بالخط الكبير على الشكل Σ .

وكما ذكرنا سابقا إذا كان لدينا مجتمع من القياسات وتباينه σ^2 غير معروف أو غير متوفر فيمكن أخذ عينة من هذا المجتمع حجمها n ، مثلا ، وحساب تباينها ثم اعتبار هذا التباين تقديرا أو تخمينا لتبـاين المجتمع الذي نجهله . ومن الطبيعي أن يكون تباين العينة ، وفقا لتعريف التبـاين ، مساويا لمتوسط مربعات انحرافـات القياسات الـ n في العينة عن متوسطها . ولكن يبرهن في الإحصاء الريـاضي أن تباين العينة سيكون تقديرا أفضل لتباين المجتمع إذا قمنـا بتعديل طفيف جدًا في صيغة التعريـف . وهذا التعديل هو أن نقسم مجموع مربعات الانحرافات عن المتوسط $\left[\sum_{i=1}^{n} (x_i - \bar{x})^2 \right]$ على $(n-1)$ بدلا

من قسمتها على (n) . وهكذا سنرمز لتباين عينة بـ S^2 ، تمييزا له عن σ^2 تباين المجتمع ، ونعرفه كما يلي :

(١ ـ ٨ ـ ٦) تعريف تباين عيّنة

تباين عينة من القياسات $x_1 , x_2 , \ldots , x_n$ هو:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

حيث $\bar{x}$ متوسط العيّنة .

(١ ـ ٨ ـ ٧) تعريف الانحراف المعياري لعيّنة

الانحراف المعياري لعيّنة من القياسات $x_1 , x_2 , \ldots , x_n$ هو الجذر التربيعي الموجب لتباين العينة .

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

وفيما تبقى من هذا الفصل سنعتبر تباين أي جملة من القياسات تباين عينة ، ونطبق لحساب التباين التعريف (١ ـ ٨ ـ ٦) وحيثما وردت كلمة التباين أو الانحراف المعياري ، فيما بقي من هذا الفصل ،فسنعني بها تباين العينة (S^2) كما عرفناه في (١ ـ ٨ ـ ٦) ، والانحراف المعياري لعينة (S) كما عرفناه في (١ ـ ٨ ـ ٧) ، إلا إذا ذكرنا ما يخالف ذلك .

مثال (١ ـ ١٩)

لتكن جملة القياسات $5 , 7 , 1 , 2 , 4$ ، احسب متوسط الانحراف ، والتباين ، والانحراف المعياري .

$$\bar{x} = \frac{19}{5} = 3.8$$

الحل

ننظم الجدول التالي بعد حساب المتوسط $\bar{x}$. ثم نطبق التعريف مباشرة لنجد :

	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
	5	1.2	1.44
	7	3.2	10.24
	1	-2.8	7.84
	2	-1.8	3.24
	4	0.2	0.04
المجموع	19	0	22.8

$$D = \frac{1}{n} \sum_{i=1}^{n} |x_i - \bar{x}| \quad \text{(متوسط الانحراف)}$$

$$= \frac{1}{5}(1.2 + 3.2 + 2.8 + 1.8 + 0.2) = \frac{9.2}{5} = 1.84$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{22.8}{4} = 4.56 \quad \text{(التباين)}$$

$$S = \sqrt{4.56} = 2.14 \quad \text{(الانحراف المعياري)}$$

والتطبيق المباشر للتعريف يتطلب جهدا حسابيا لا مسوّغ له . (يتضمن $2n + 4$ عملية حسابية) ويعاني، في الغالب، من نقص في الدقة . وإذ نحسب قبل كل شيء المتوسط $\bar{x}$ ، نبدأ بعملية تقسيم، وإذا كانت عملية التقسيم غير منتهية فسيؤثر ذلك على دقة النتائج اللاحقة . وسنقدم الآن صيغة مختزلة لحساب التباين تختصر الجهود الحسابية وتعطي التباين بدقة أكبر.

(١ ـ ٨ ـ ٨) صيغة مختزلة لحساب التباين

باستخدام خواص الرمز Σ المذكورة في البند (٥) من الملحق (١)، ومن تعريف المتوسط يمكن أن نكتب ما يلي :

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}\left(x_i^2-2\bar{x}x_i+\bar{x}^2\right)$$

$$=\sum_{i=1}^{n}x_i^2-\sum_{i=1}^{n}2\bar{x}x_i+\sum_{i=1}^{n}\bar{x}^2$$

$$=\sum_{i=1}^{n}x_i^2-2\bar{x}\sum_{i=1}^{n}x_i+n\bar{x}^2$$

$$=\sum_{i=1}^{n}x_i^2-2\bar{x}\,(n\bar{x})+n\bar{x}^2$$

$$=\sum_{i=1}^{n}x_i^2-n\bar{x}^2=\sum_{i=1}^{n}x_i^2-n\left(\frac{\sum_{i=1}^{n}x_i}{n}\right)^2$$

$$=\sum_{i=1}^{n}x_i^2-\frac{\left(\sum x_i\right)^2}{n}$$

ومنه :

$$S^2=\frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2-\frac{\left(\sum x_i\right)^2}{n}\right]$$

ولم نعد نستهل العمل الحسابي بعملية تقسيم، بالإضافة إلى أن هذه العبارة تتضمن $(n+6)$ عملية حسابية مما يوفر $(n-2)$ عملية حسابية. وهي بذلك أسرع وأدق من التطبيق المباشر للتعريف.

ونكتب العبارة المختزلة السابقة، أحيانا، على الشكل :

$$S^2=\frac{1}{n(n-1)}\left[n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2\right]$$

وهي تتضمن عملية تقسيم واحـدة تشكل خاتمة العمل الحسابي. ومع مـا يبدو للوهلة الأولى من تعقيـد في كتابة الصيغة المختـزلة، إلا أن كل ما نحتاجـه لتطبيقها هو مجموع القياسات ومجموع مربعاتها وعددها.

مثال (١ ـ ٢٠)

احسب تباين القياسات في المثال (١ ـ ٦) بتطبيق الصيغة المختزلة.

الحل

ننظم الجدول المبين جانبا ثم نطبق الصيغة المختزلة فنجد:

	x_i	x_i^2
	5	25
	7	49
	1	1
	2	4
	4	16
المجموع	19	95

$$S^2 = \frac{1}{4}\left[95 - \frac{(19)^2}{5}\right]$$

$$= \frac{1}{4}\left(95 - \frac{361}{5}\right) = 4.56$$

(١ ـ ٨ ـ ٩) حساب التباين في بيانات مصنفة

لنعد إلى الفقرة (١ ــ ٥)، وبخاصة إلى الجدولين (١ ـ ٦) و (١ ــ ٧)، ولنحاول تطبيق التعريف (١ ـ ٨ ـ ٦) فالمطلوب إذا هو حساب انحراف كل قياس y_i عن المتوسط $\bar{y}$، ثم أخذ مجموع مربعات هذه الانحرافات. وإذا كان القياس y_1، مثلا، مكررا f_1 مرة، فسيتضمن مجموع مربعات الانحرافات حدودا متطابقة ومكررة مثل:

$$\underbrace{(y_1-\bar{y})^2+(y_1-\bar{y})^2+\cdots+(y_1-\bar{y})^2}_{\text{مكرر } f_1 \text{ مرة}}$$

ومن الأفضل بـالطبع ، كتابـة مجمـوع حدود مطـابقـة لبعضها مثل هـذه الحدود ، على الشكل

$$f_1\,(y_1-\bar{y})^2$$

والأمر نفسه في بقيـة الحدود ، وهكذا تتخـذ العلاقـة الواردة في تعـريف تبايـن العينة ، الصيغة التالية من أجل بيان مرتب :

$$S^2=\frac{1}{\sum_1^m f_i-1}\left[\sum_1^m f_i(y_i-\bar{y})^2\right]$$

حيث :

$$\bar{y}=\frac{\sum_1^m f_i\,y_i}{\sum_1^m f_i}$$

وتصبح الصيغة المختزلة لحساب التباين كما يلي :

$$S^2=\frac{1}{n-1}\left[\sum_1^m f_i\,y_i^2-\frac{\left(\sum_1^m f_i\,y_i\right)^2}{n}\right],$$

حيث $n=\sum_1^m f_i$.

مثال (١ ـ ٢١)

قذفنا حجر نرد مائة مرة فكانت تكرارات النتائج الست الممكنة كما يلي :

جدول (١ ـ ١٦)

y_i	1	2	3	4	5	6
f_i	19	15	15	20	14	17

والمطلوب حساب تباين هذا التوزيع التكراري وانحرافه المعياري.

الحل

لحساب التباين ننظم الجدول التالي

جدول (١ ـ ١٧)

	y_i	f_i	$f_i y_i$	$f_i y_i^2$
	1	19	19	19
	2	15	30	60
	3	15	45	135
	4	20	80	320
	5	14	70	350
	6	17	102	612
المجموع		$100 = \Sigma f_i$	$346 = \Sigma f_i y_i$	$1496 = \Sigma f_i y_i^2$

والتباين المطلوب (S^2) هو:

$$S^2 = \frac{1}{99}\left[1496 - \frac{(346)^2}{100}\right] = 3.02$$

والانحراف المعياري (S) هو

$$S = \sqrt{3.02} = 1.74$$

ولحساب تباين بيان مصنف (أو مبوب) نعتبر أن جميع القياسات التي تنتمي إلى فئة مساوياً لمركز هذه الفئة والخطوات الحسابية هي بالضبط كما في حالة بيان الرتب، حيث y_i هي الآن مركز الفئة، و f_i التكرار الموافق لهذه الفئة. وللتوضيح نأخذ المثال التالي.

مثال (١ ـ ٢٢)

احسب التباين والإنحراف المعياري للتوزيع التكراري في الجدول (١ ـ ٦).

الحل

ننظم الجدول التالي:

جدول (١ ـ ١٨)

	y_i مركز الفئة	f_i التكرار	$f_i y_i$	$f_i y_i^2$
	84	1	84	7056
	89	2	178	15842
	94	4	376	35344
	99	7	693	68607
	104	9	936	97344
	109	10	1090	118810
	114	7	798	90972
	119	6	714	84966
	124	4	496	61504
المجموع		$50 = \Sigma f_i$	$5365 = \Sigma f_i y_i$	$580445 = \Sigma f_i y_i^2$

ويكون التباين المطلوب (S^2) هو

$$S^2 = \frac{1}{49}\left[580445 - \frac{(5365)^2}{50}\right]$$
$$= \frac{1}{49}[580445 - 575664.5]$$
$$= \frac{4780.5}{49} = 97.56$$

والانحراف المعياري (S) هو

$$S = \sqrt{97.56} = 9.88$$

والجدير بالـذكر أننا لو حسبنا الإنحـراف المعياري من البيان الأصلي المعطى في الجدول (١ ـ ٤) مباشرة لحصلنا على $S = 10.05$.

(١ ـ ٨ ـ ١٠) أثر الانسحاب وتغيير سلم القياس في التباين

لتكـن $x_1, x_2, ..., x_n$ جملـة من القيـاسات، متوسطهـا $\bar{x}$ ، وتباينها S_x^2 . إذا أضفنا العدد نفسه، b مثـلا، إلى كل قياس، فينبغي ألا يـؤثر ذلك على التبـاين . وإذ تعتمـد قيمة التبـاين على الفروق بين القيـاسـات، فإن الفرق بين أي قيـاسين لن يتغير

عندما نضيف إلى كل منهما العدد نفسه [انظر البند (٨) من الملحق ١]. أما إذا ضربنا كل قياس بعدد، a مثلا، فسيُضرب التباين بمربع هذا العدد، a^2، ويضرب الانحراف المعياري بالقيمة المطلقة للعدد a. ويمكن بيان ذلك في المحاكمة البسيطة التالية:

لنرمز بـ y_i للقياس الناتج عن ضرب x_i بـ a ثم إضافة b إلى الناتج، أي لنفرض أن:

$$y_i = ax_i + b \;\; ; \; i = 1, 2, \ldots, n$$

فنعلم من خواص المتوسط (الفقرة ١ ـ ٧ ـ ٢) أن:

$$\bar{y} = a\bar{x} + b$$

حيث يرمز $\bar{y}$ للمتوسط الجديد. ومن تعريف التباين وخواص المجموع Σ [انظر البند (٥) من الملحق (١)] يمكن أن نكتب:

$$S_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n-1}\left[(ax_i + b) - (a\bar{x} + b)\right]^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left[a(x_i - \bar{x})\right]^2 = a^2 \times \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = a^2 S_x^2$$

وبأخذ الجذر التربيعي الموجب للطرفين نجد:

$$S_y = |a|\, S_x$$

وهذا يعني أن عملية الانسحاب لا تؤثر في التباين كما توقعنا، ولكن عملية تغيير سلم القياس لها أثر كبير في التباين. وعلى سبيل المثال، إذا كانت القياسات x_i مقاسة بالسنتيمتر، وغيرنا وحدة القياس إلى الميلليمتر، أي ضربنا كل قياس بـ 10، فإن التباين سيضرب بمائة، وسيضرب الانحراف المعياري بعشرة.

(١ ـ ٩) حساب المتوسط والانحراف المعياري من خلال تحويل البيان الإحصائي

سنقدم فيما يلي طريقة لحساب المتوسط والانحراف المعياري توفر الكثير من الجهود الحسابية، وذلك في حالة بيان مصنف أطوال الفئات فيه متساوية. وهي طريقة عامة وسهلة التطبيق، فلنفرض أن طول الفئة ω، وأن y_0 مركز الفئة الواقعة في الوسط تماما إذا كان عدد الفئات فرديا، أو مركز إحدى الفئتين الواقعتين في الوسط إذا كان عدد الفئات زوجيا. ولنطبق على مراكز الفئات التحويل:

$$Z_i = \frac{y_i - y_0}{\omega}$$

أي نطرح من مركـز كل فئة العدد y_0 ثم نقسم النـاتج على ω. ومن علاقة التحويل نستنتج أن :

$$y_i = \omega Z_i + y_0$$

وكما نعلم فإن :

$$\bar{y} = \omega \bar{Z} + y_0$$

و

$$S_y^2 = \omega^2 S_z^2 , S_y = |\omega| S_z = \omega S_z$$

ω هنا موجبة دوما باعتبارها طول فئة .

وسنجد أن المقاديـر Z_i أعداد صحيحـة متناظـرة حول الصفـر. وفي حالـة تسع فئات ، مثلا ، سنجد المقادير Z_i على الشكل $-4, -3, -2, -1, 0, 1, 2, 3, 4$ وبالطبع فإن التعامل مع هذه الأعداد أسهل كثيرا . والآن نعتبر هذه الأعداد مراكز للفئات وننجز الحسابـات تماما كما في الفقـرة السابقـة (١ ـ ٨ ـ ٣) فنحصل على $\bar{Z}$ و S_z^2 و S_Z بسهـولـة ، ومنها نستنتـج المتوسط والتبـاين والانحـراف المعيـاري للبيـان الأصلي قبل التحويل من خلال العلاقات :

$$\bar{y} = \omega \bar{Z} + y_0 \ , S_y^2 = \omega^2 S_z , S_y = \omega S_z$$

مثال (١ ـ ٢٣)

بالعـودة إلى المثـال (١ ـ ٢٢) ، احسب المتـوسط والانحـراف المعيـاري بطريقـة تحويل البيان الإحصائي .

الحل

لدينا تسع فئات ، والفئة الواقعة في الوسط هي الفئة الخامسة ومركزها $y_0 = 104$. وطول الفئة $\omega = 5$. وبإجراء التحويل :

$$Z_i = \frac{y_i - 104}{5}$$

تصبح مراكز الفئات

$$Z_1 = \frac{y_1 - 104}{5} = \frac{84 - 104}{5} = -4$$

$$Z_2 = \frac{y_2 - 104}{5} = \frac{89 - 104}{5} = -3$$

وهكذا.

وبدلا من الجدول (١ ـ ١٨) ننظم الجدول (١ ـ ١٩)، التالي :

جدول (١ ـ ١٩)

مركز الفئة yi	التكرار f_i	Z_i	$f_i Z_i$	$f_i Z_i^2$
84	1	-4	-4	16
89	2	-3	-6	18
54	4	-2	-8	16
99	7	-1	-7	7
104	9	0	0	0
109	10	1	10	10
114	7	2	14	28
119	6	3	18	54
124	4	4	16	64
المجموع	50		33	213

$$\bar{Z} = \frac{33}{50} = 0.66$$

$$S_z^2 = \frac{1}{49}\left[213 - \frac{(33)^2}{50}\right] = \frac{191.22}{49}$$

ومنه

$$\bar{y} = \omega \bar{Z} + 104 = 5 \times 0.66 + 104 = 107.3$$

$$S_y^2 = \omega^2 S_z^2 = 25 \times \frac{191.22}{49} = 97.56 \; ; \; S_y = 9.88$$

وهي الأجوبة ذاتها التي حصلنا عليها في المثال (١ ـ ٢٢).

(١ ـ ١٠) حول الأهمية العملية للمتوسط والانحراف المعياري

من الطبيعي أن نتساءل عن مدى نجاح التباين S^2 في التعبير عن خاصية التغير في جملة من القياسات. وسنجد الجواب الصريح عن هذا التساؤل في نقطتين نعرضهما فيما يلي:

١ ـ لنأخذ مجموعة القياسات $1, 2, 3, 4$ ، ولنحسب تباينها:

$$S^2 = \frac{1}{3}\left[\left(1^2 + 2^2 + 3^2 + 4^2\right) - \frac{(1+2+3+4)^2}{4}\right]$$
$$= \frac{1}{3}[30 - 25] = \frac{5}{3}$$

ولنحسب، على الوجه الآخر، الفروق بين كل قياس والقياسات الباقية، كما في الجدول (١ ـ ٢٠)، ثم لنحسب متوسط مربعات هذه الفروق فنجد $\frac{40}{12} = \frac{10}{3}$. أي أن متوسط مربعات الفروق الموجودة بين القياسات كافة يساوي $2S^2$.

جدول ١ ـ ٢٠

	1	2	3	4
1	0	-1	-2	-3
2	1	0	-1	-2
3	2	1	0	-1
4	3	2	1	0

*وبصورة عامة

إذا كانت $x_n, \ldots, x_1$ عينة من القياسات، متوسطها $\bar{x}$ وتباينها S^2 ، فإن عدد الأزواج المختلفة من القياسات التي يمكن تشكيلها هو $n(n-1)$. ومتوسط مربعات الفروق بين العددين في كل زوج منها هو:

*للقراءة فقط.

$$\frac{1}{n(n-1)}\sum_{i}^{n}\sum_{j}^{n}(x_i - x_j)^2$$

$$= \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(x_i - \bar{x}) - (x_j - \bar{x})\right]^2$$

$$= \frac{1}{n(n-1)}\sum_{i}^{n}\sum_{j}^{n}\left[(x_i - \bar{x})^2 + (x_j - \bar{x})^2 - 2(x_i - \bar{x})(x_j - \bar{x})\right]$$

$$= \frac{1}{n(n-1)}\sum_{i}^{n}\left[n\ (x_i - \bar{x})^2 + (n-1)\,S^2 + 0\right]$$

$$= \frac{1}{n(n-1)}\left[n(n-1)\ S^2 + n(n-1)\ S^2\right] = 2S^2.$$

وهـذا يوضح أن التبـاين S^2 يلخص بأمـانـة كافـة التغيرات من قيـاس إلى آخـر التي يتضمنها البيان الإحصائي. وبالتالي فإنه يشكل تعبيرا ناجحا عن خاصية التغير ضمن البيان الإحصائي.

٢ ـ هنــاك متباينـة مشهـورة تسمى متباينـة تشيبيشيف، ويمكن التعبير عنها بطـريقة مبسطة كما يلي:

لتكن $x_1, x_2, ..., x_n$ جملة من القيـاسات متوسطها $\bar{x}$ وانحـرافها المعياري S. وليكن t عددا أكبر من الواحد أو يساويه، فـالنسبة من هذه القياسات التي تقع ضمن الفترة $(\bar{x} - ts\ , \bar{x} + ts)$ لا تقل عن $1 - \frac{1}{t^2}$.

لنختر الآن بعض القيم لـ t، ولنحسب النسبة $1 - \frac{1}{t^2}$ فنجد:

جدول (١ ـ ٢١)

t	1	2	3
$1 - \frac{1}{t^2}$	0	$\frac{3}{4}$	$\frac{8}{9}$

فالمتباينة لا تقدم أية معلومات من أجل $t = 1$. ولكنها تقول، في حالة $t = 2$ ؛ أن ثلاثة أربـاع القياسـات، على الأقل، واقع ضمن فترة تمتـد ضعف الانحراف المعيـاري على جانبي المتوسط. أي تقع ضمن الفترة $(\bar{x} - 2S\,,\ \bar{x} + 2S)$ وتقول في حالة $t = 3$ ، أن ما لا يقل عن ثمانية أتساع القياسات (89% تقريبا) واقع ضمن فترة تمتـد بمقدار ثلاثة انحرافات معيارية على جانبي المتوسط، أي تقع بين العدد $\bar{x} + 3S$ والعدد $\bar{x} - 3S$.

مثال (١ ـ ٢٤)

لنعد إلى البيان الإحصـائي في الجدول (١ ـ ٤)، فقــد حسبنا في المثـال (١ ـ ٦) متـوسطـه فوجـدنـاه $\bar{x} = 107.3$ وحسبنـا في المثـال (١ ـ ٢٢) الانحـراف المعيـاري فوجدناه $S = 9.88$. ولدينا

$$\bar{x} - 2S = 107.3 - 2 \times 9.88 = 87.54$$

$$\bar{x} + 2S = 107.3 + 2 \times 9.88 = 127.06$$

ولو تفقدنا القياسات الخمسين في الجدول (١ ـ ٥)، لوجدنا أن 49 منها واقع بين 87.54 و 127.06 ، وهي تشكل نسبة $\frac{49}{50} = 98\%$ من القياسات.

ومما تقدم نستنتج بوضوح أن التباين S^2 يشكل مقياسا كميـا ناجحا تماما للتعبير عن خاصية التغير ضمن بيان إحصائي. وأصبح واضحا الآن أن متوسط بيان إحصائي $\bar{x}$ ، وانحرافه المعياري S ، يلخصان بصورة جيدة قياسات ذلك البيان. ومن خلالهما، يمكن تشكيل صـورة ذهنية جيـدة للغايـة عن التـوزيع التكراري للبيـان دون أن نعلم مفردات البيان.

وعلى سبيل المثال، لـو قيل لنـا أن درجـات فصل يتألف من 40 طالبـا في مادة الإحصاء، لها متـوسط يساوي 72 ، وانحراف معيـاري يساوي 8 ، لأمكننا باستخدام هذين الـرقمين فقط، تقديم الوصف التالي لتـوزيع الدرجات، دون أن تكون لـدينا أية معلومات أخرى عن واقع الدرجات نفسها:

تتمركز الدرجات في هـذا الفصل حول القيمة 72 ، وما لا يقل عن ثـلاثين طالبا حصلوا على درجات تتراوح بين $72 - 2 \times 8 = 56$ و $72 + 2 \times 8 = 88$. وما لا يقل عن

36 طالبا من الطلاب الأربعين حصلوا على درجات تتراوح بين $48 = 8 \times 3 - 72$ و $96 = 8 \times 3 + 72$.

ويجدر الانتباه إلى عبارة «ما لا يقل» فمتباينة تشيبيشيف متحفظة، وفي معظم الحالات تكون النسبة الفعلية أكبر من $1 - \frac{1}{t^2}$ خاصة إذا كان البيان الإحصائي قريبا من التناظر.

(١ ـ ١١) معامل التغير

رأينا أن التباين يعبر بنجاح عن خاصية التغير في بيان إحصائي. ومن الطبيعي أن يكون البيان الإحصائي أكثر تجانسا كلما كانت قياساته أقل تغيرا من أحدها إلى الآخر. وكلما زاد التباين استنتجنا أن البيان الإحصائي أقل تجانسا، ولكن هب أننا نريد مقارنة بيانين إحصائيين من حيث أيهما أكثر تجانسا من الآخر، فهل يمكن الاعتماد على مقارنة تباينيهما وإعطاء حكم في هذه المسألة؟ لقد وجدنا في الفقرة (١ ـ ٨ ـ ٣) أن مقدار التباين يعتمد على وحدة القياس المستخدمة في البيان الإحصائي مما يجعله غير صالح للمقارنة بين عينتين من القياسات من حيث درجة التجانس في كل منهما. وهناك عامل آخر، إذ بالرغم من استخدام وحدة القياس نفسها في البيانين اللذين نريد مقارنتهما، إلا أن طبائع الأمور قد تجعل أرقام البيان الأول كبيرة، وأرقام البيان الثاني صغيرة. كأن يتضمن البيان الأول أوزان مجموعة من العجول بالكيلوغرام، ويتضمن البيان الثاني أوزان مجموعة من الفراريج بالكيلوغرام أيضا. ونوضح بالمثال التالي:

مثال (١ ـ ٢٥)

في مزرعة خمسة عجول، وعشرون فروجا، سجلنا الأوزان ضمن كل مجموعة بالكيلوغرام فحصلنا على البيانين التاليين:

العجول : 285.50; 280.40; 283.00; 280.75; 281.40;

الفراريج : 1.50; 1.40; 0.95; 1.35; 1.45; 1.05; 1.05;
0.99; 1.45; 1.50; 1.35; 1.45; 1.00; 1.10;
1.25; 1.35; 1.10; 1.45; 1.00; 1.20;

احسب المتوسط والتباين والانحراف المعياري لكل منها .

الحل

	المتوسط	التباين	الانحراف المعياري
العجول	282.21	4.37	2.09
الفراريج	1.26	0.036	0.19

ولو استخدمنا ، في المثال السابق ، الانحراف المعياري للمقارنة والحكم على درجة تجانس كل من البيانين ، لاستنتجنا خطأ أن مجموعة الفراريج أكثر تجانسا من مجموعة العجول ، لأن انحرافها المعياري ، وبالتالي تباينها ، أصغر بكثير. ولكن صغر الانحراف المعياري للفراريج ، يعود إلى صغر أوزان الفراريج بالمقارنة مع أوزان العجول ، وليس لكونها أكثر تجانسا .

وسنعرف الآن مقياسا يسمى معامل التغير، وهو لا يعتمد على وحدة القياس المستخدمة ، ولا يتأثر بكون القياسات كبيرة أو صغيرة ، مما يجعله صالحا لمقارنة درجتي التجانس في عينتين من القياسات ، وذلك بصرف النظر عن طبيعة هذه القياسات أو عن وحدات القياس المستخدمة في كل منهما .

تعريف معامل التغير

معامل التغير، ونرمز له بـ $c.v$ ، لجملة من القياسات متوسطها $\bar{x}$ ، وانحرافها المعياري S ، هو بالتعريف :

$$c\,.\,v = \frac{S}{\bar{x}}$$

ونعلم من خواص المتوسط وخواص التباين أنه إذا ضربنا كل قياس في جملة من القياسات بعدد معين ، فإن كلا من المتوسط $\bar{x}$ والإنحراف المعياري S ، سيضرب

بالعدد نفسه، (أو يقسم على العدد نفسه) وبالتالي ستبقى النسبة $\frac{S}{\bar{x}}$ بدون تغيير. وإذا كانت أرقام أحد البيانين كبيرة بطبيعتها وأرقام الآخر صغيرة، فإن قسمة S على $\bar{x}$ يعطينا الانحراف المعياري لكل وحدة قياس، مما يخلص معامل التغير من أي أثر لحجم القياسات.

مثال (١ ـ ٢٦)

في المثال السابق (١ ـ ٢٥)، أحسب معامل التغير لكل من جملتي القياسات وقارنهما من حيث درجة التجانس ضمن كل منهما.

الحل

$$c.v.\ (\text{للعجول}) = \frac{S}{\bar{x}} = \frac{2.09}{282.21} = 0.007$$

$$c.v.\ (\text{للفراريج}) = \frac{S}{\bar{x}} = \frac{0.19}{1.26} = 0.15$$

ويتضح الآن أن مجموعة العجول أكثر تجانسا بكثير من مجموعة الفراريج، فمعامل تغيرها *0.007* ، بينما معامل تغير الفراريج *0.15* ، وهو أكبر من معامل تغير العجول بما ينوف على إحدى وعشرين مرة.

(١ ـ ١٢) القيمة المعيارية

وسننتقل الآن إلى مشكلة أخرى، فلنفرض أن لدينا جملتين من القياسات، فكيف يمكن، عند الحاجة، مقارنة قياس من الجملة الأولى بقياس من الجملة الثانية؟

وعلى سبيل المثال

لنفرض أن الجملتين من القياسات هما درجات طلاب الفصل في مادة الرياضيات، ودرجاتهم في مادة اللغة العربية، ونريد مقارنة درجتي طالب معين في المادتين، وإذا فرضنا أن درجته في الرياضيات كانت *70* ، وأنها في اللغة العربية *60* ،

فهل يعني ذلك أن تحصيله في الرياضيات أفضل من تحصيلـه في اللغة العربية؟ المسألة هنا نسبية، فقد يكون معظم طـلاب الفصل نالوا درجات أعلى من 70 في الرياضيات، ولكن قليلا منهم فقط نـال درجات تزيـد على الستين في اللغة العربيـة. وفي مثل هذه الحالة تنعكس الآية فنقول، على عكس ما يوحيه الرقمان، إنه كان من المتفوقين في اللغة العـربية، ومن المقصرين في الـريـاضيات. والطـريقة التي تسمـح لنا بمـراعاة الـواقع النسبي، واتخاذ الحكم الصحيح، هي حساب متـوسط كل جملة وانحرافهـا المعياري. ثم نرد كل درجة إلى ما يسمى بقيمتها المعيارية، بأن نطرح منها المتوسط ثم نقسم الناتج على الانحراف المعياري. وبذلك نحسب كم انحرافا معياريا تبتعد الدرجة عن متوسط الـدرجات؟ أو بعبـارة أخرى، نقيس الفـرق بين الدرجـة والمتوسط بـوحدة قيـاس هي الانحـراف المعياري للـدرجـات. ولنفرض في مثـالنا هنـا أن متـوسط درجات طـلاب الفصـل في مـادة الرياضيـات كـان 75 بانحـراف معياري يسـاوي 5 ، وأن متـوسط درجـات الطلاب في مادة اللغـة العربيـة كان 52 ، بانحـراف معياري يسـاوي 6 .
فالـدرجة المعياريـة في الرياضيات هي $1- = \frac{70-75}{5}$ والـدر جة المعياريـة في اللغة العـربيـــة هي $1.33 = \frac{60-52}{6}$. وهي أكبر من 1- ، أي أن تحصيله في اللغة العربية أفضل.

تعريف القيمة المعيارية

إذا كـان $\bar{x}$ و S متوسط جملـة من القياسات وانحرافهـا المعياري، على الترتيب. فنعرف القيمة المعيارية لأي قياس، x ، من هذه الجملة، بأنها:

$$\frac{x-\bar{x}}{S}$$

ونلاحظ أن رد جملـة من القيـاسـات إلى الشكل المعيـاري، أو معـايــرة جملـة القياسـات، يجعل متـوسطها مسـاويا للصفـر، وانحرافهـا المعياري مسـاويا لـلواحد الصحيح. ولبيان ذلك نكتب:

لتكن $x_1, x_2, \ldots, x_n$ جملة من القياسـات متوسطها $\bar{x}$ وانحرافها المعياري S . ووفقا لتعريف المعايرة يمكن أن نكتب:

$$Z_i = \frac{x_i - \bar{x}}{S} \quad , \quad i = 1, 2, \ldots, n.$$

حيث رمزنا بـ Z_i للقيم المعيارية. لنحسب الآن: $\bar{Z}$ متوسط القيم المعيارية و S_z^2 انحرافها المعياري فنجد:

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^{n} Z_i = \frac{1}{n} \sum_{i=1}^{n} \frac{x_i - \bar{x}}{S} = \frac{1}{ns} \sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

$$S_z^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Z_i - \bar{Z})^2 = \frac{1}{n-1} \sum_{i=1}^{n} Z_i^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{S} \right)^2$$

$$= \frac{1}{S^2(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{S^2}{S^2} = 1$$

وهذا يعني أن معايرة جملتين من القياسات تردهما إلى جملتين لهما المتوسط نفسه، وهو الصفر، والانحراف المعياري نفسه، وهو الواحد.

بقيت ملاحظة أخيرة، وهي أنه إذا كانت $Z_1, Z_2, \ldots, Z_n$ القيم المعيارية لجملة من القياسات فإن

$$\sum_{i=1}^{n} (Z_i - \bar{Z})^2 = (n-1)\, S_z^2 = (n-1) \times 1 = n - 1$$

وبما أن $\bar{Z} = 0$ ، فإن

$$\sum_{i=1}^{n} Z_i^2 = n - 1$$

أي أن مجموع مربعات القيم المعيارية لجملة من القياسات يساوي عدد القياسات n مطروحا منه الواحد. وسنستفيد من هذه الخاصة في الفقرة القادمة.

تمارين (١ ـ ٤)

١) فيما يلي الأطوال بالسنتيمتر لعشرة أوراق من نبات منزلي:

10.0; 10.2; 6.5, 7.0, 7.8, 10.8, 6.1, 5.9, 8.9, 10.0

احسب المدى، ومتوسط الانحراف، والتباين، والانحراف المعياري.

٢) استخدمنا سبعة موازين حرارة لقياس درجة حرارة جسم بالتدريج المئوي. فكانت النتائج كما يلي:

$$-4.12, -4.09, -4.10, -4.08, -4.09, -4.13, -4.10$$

احسب التباين والانحراف المعياري.

٣) ماذا يمكن القول عن مجموعة قياسات تباينها يساوي الصفر؟ وإذا حسبت تباين جملة من القياسات فوجدته سالبا فماذا تستنتج؟

٤) في كل مما يلي أحسب المدى والانحراف المعياري:

أ ـ $4, 2, 8, 1, 4, 5, 8, 10, 3$

ب ـ $-5, 3, -1, -4, 3, -8, -2$

جـ ـ $1, 2, 3, 0, -3, 3, 3, -1$

تحقق في (أ) أنك إذا أخذت متوسط مربعات انحرافات كل قياس عن بقية القياسات فإن النتائج يساوي ضعفي التباين.

٥) فيما يلي التوزيع التكراري لعدد القطع المعيبة التي وجدت في 404 صناديق من القطع المصنعة.

عدد القطع المعيبة	0	1	2	3	4	5	6	7	8	9	10	11	12
عدد الصناديق	53	110	81	58	35	20	18	12	9	3	1	2	1

ا ـ احسب المتوسط والتباين ومعامل التغير.

ب ـ احسب الوسيط والمنوال.

٦) إذا كان تباين عينة تتضمن مائة قياس هو 15 ، فاحسب مجموع مربعات انحرافات القياسات عن متوسطها.

٧) تباين عينة من القياسات يساوي 20 . كم يصبح التباين :
ا - إذا ضربنا كل قياس بـ 5 ؟
ب - إذا قسمنا كل قياس على 5 ؟

٨) أخذنا عينتين من مجتمعين فأعطتا النتائج التالية :

العينة الثانية	العينة الأولى
$\sum_{i=1}^{50} y_i = 400$	$\sum_{i=1}^{30} x_i = 270$
$\sum_{1}^{50} y_i^2 = 3984$	$\sum_{1}^{30} x_i^2 = 2691$

ا - احسب تباين كل عينة .
ب - أيهما أكثر تجانسا؟
جـ - إذا دمجنا العينتين في عينة واحـدة فاحسب متوسط العينة الجديـدة ومعـامل تغيرها .

٩) احسب نصف المدى الربيعي ومعـامل التغير في كل من التمارين ٤ ، ٥ ، ٦ ، ٧ من مجموعة التمارين (١ - ١) .

١٠) احسب التباين في كل من التمرينين ٩ ، ١١ من مجموعة التمارين (١ - ٢) .

١١) احسب الانحراف المعياري ومعامل التغير لسماكة الجلـد في التمرين ٨ من مجموعة التمارين (١ - ١) ثم تحقق من أن 95% تقريبا من القياسات واقع في حدود انحرافين معياريين عن يمين ويسار المتوسط .

١٢) بالإشارة إلى التمرين ٧ من مجموعـة التمارين (١ - ٣)، احسب التباين والانحراف المعياري لنصف قطر رد الفعل لاختبار الليبرومين في كل من التوزيعات الثلاثة .

١٣) احسب الانحراف المعياري ومعامل التغير لكل من البيانات المعطاة في التمارين ٦، ٧، ١٠ من مجموعة التمارين (١ ـ ١).

١٤) فيما يلي بيانات تتعلق بمنطقة معينة لعامي ١٩٥٩ و١٩٦٠. أحسب لكل بيان، المدى، والانحراف المعياري، ومعامل التغير. أي البيانات الثلاثة أكثر تجانسا؟

متوسط الرطوبة النسبية عند التاسعة صباحا (٪)	متوسط درجة الحرارة (بالفهرنهايت)	معدل سقوط المطر (بالبوصة)	الشهر
78	72.1	1.45	يناير
78	72.5	1.44	فبراير
78	72.1	2.69	مارس
77	72.6	5.15	أبريل
79	73.3	7.46	مايو
85	73.2	0.73	يونيو
72	72.8	0.51	يوليو
78	71.9	5.17	أغسطس
78	71.4	4.20	سبتمبر
78	71.7	4.08	أكتوبر
78	71.6	6.68	نوفمبر
79	71.6	2.77	ديسمبر

١٥) في تجربة لتقدير فائدة مضاد للتسمم في معالجة الكزاز، قورنت مجموعة تناولت المضاد مع مجموعة لم تتناوله. وقد تم تخصيص المرضى للمجموعتين بطريقة عشوائية، وفيما يلي بيان بأعمار المرضى. ارسم مضلع التكرار النسبي المتجمع الصاعد لكل من المجموعتين على حدة واستخدمها لتقدير العمر الوسيط ونصف المدى الربيعي لكل مجموعة.

١٦) احسب نصف المدى الربيعي ومعامل التغير في التمرين (٥) من مجموعة التمارين (١ ـ ٣).

لم يتناول مضاد للتسمم (N)		تناول مضاد للتسمم (A)	
33	18	16	41
20	24	28	28
39	19	27	35
36	12	20	40
30	29	17	30
60	14	12	27
17	18	12	50
27	18	16	30
33	50	20	9
14	16	10	40
10	14	11	30
60	52	20	18
12	16	50	31
24	40	29	14
12	30	24	25
10	40	14	27
60	40	17	16
27	27	25	36
8	20	10	25
		24	40
			22

١٧) إذا علمت توزيع التكرار النسبي التالي :

الفئة	19.5 - 39.5	39.5 - 59.5	59.5 - 79.5	79.5 - 99.5
التكرار النسبي	0.12	0.28	0.36	0.24

فاحسب المتوسط ، الوسيط ، المنوال ، Q_1 ، Q_3 ، σ^2 .

١٨) كان متوسط معدلات الطلبة المتقدمين لإحدى الجامعات 20.4 بانحراف معياري 3.1 ، ومتوسط معدلات الطلبة المتقدمين لجامعة أخرى 21.1 بانحراف معياري 2.8 .
إذا تقدم طالب معدله 25 إلى كل من الجامعتين ففي أيهما ستكون فرصة قبوله أفضل؟

١٩) في دراسة قام بها مركز للأغذية تبين أن متوسط مقدار الفيتامين B في عدد من شرائح الخبز هو 0.26 ملغم. بانحراف معياري قدره 0.005 ملغم. استخدم هذه المعلومات لإكمال العبارات التالية :
- ما لا يقل عن 25/36 من هذه الشرائح يحتوي على مقدار من فيتامين B واقع بين (. . . و . . .).
- ما لا يقل عن 63/64 من هذه الشرائح يحتوي على مقدار من فيتامين B واقع بين (. . . و . . .).

٢٠) إذا علمت أن معامل تغير بيان إحصائي يتضمن ثمانين قياسا هو 0.1 وأن مجموع قياساته 800، فاحسب مجموع مربعات القياسات $\sum x_i^2$.

٢١) قمنا بدراسة زمنية لتحديد الوقت الذي يستغرقه إنجاز عملية معينة في مؤسسة صحية. وقد قسنا الزمن الضروري لإنجاز هذه العملية لكل من 40 عاملا، ووجدنا أن المتوسط يساوي 12.8 وحدة زمن بانحراف معياري يساوي 1.7 وحدة زمن. والمطلوب إعطاء وصف للبيان الإحصائي مستخدما متباينة تشيبيشيف.

٢٢) لديك المعلومات التالية عن أسعار مجموعة من مطاعم الدرجة الأولى في مدينة معينة :

الوجبة	متوسط الكلفة	الانحراف المعياري S
لحم	24.25 ر. س	1.50
دجاج	13.72 ر. س	0.94
سمك	33.65 ر. س	1.13

وأحد هذه المطاعم ويسمى «مطعم التوفير» يقدم وجبة اللحم في مقابل 28 ريالا، ووجبة الدجاج في مقابل 17 ريالا، ووجبة السمك في مقابل 36 ريالا، **هل تعتقد** أن هذا المطعم يستحق الإسم الذي يدّعيه؟ ولماذا؟

(١ ـ ١٣) الارتباط

(١ ـ ١٣ ـ ١) مقدمة

لدينا مجموعة n من الأشخاص، ولنفرض أننا قمنا بقياس ظاهرتين لدى كل شخص منها، ورمزنا لقياس إحداهما بـx، ولقياس الأخرى بـy (مثلا، x ترمز للطول، y ترمز للوزن). فحصلنا بذلك على n من أزواج الأعداد، (x_1, y_1) لأول شخص، (x_2, y_2) للشخص الثاني، . . . ، وأخيرا (x_n, y_n) للشخص الأخير.

لنرتب القيم x_i من الأصغر إلى الأكبر، ثم لنضع أمام كل قيمة لـx قيمة y الموافقة لها. ولنفرض أننا وجدنا قيم y مرتبة أيضا من الأصغر إلى الأكبر، فأصغر قيمة لـx قابلتها أصغر قيمة لـy (أي أن الشخص ذا الطول الأصغر كان أيضا ذا الوزن الأصغر بين الأشخاص الـn الخاضعين للتجربة)، والقيمة بعد الصغرى لـx قابلتها القيمة بعد الصغرى لـy، . . . . ، وأخيرا مقابل أكبر قيمة لـx كان بين قيم y أكبرها.

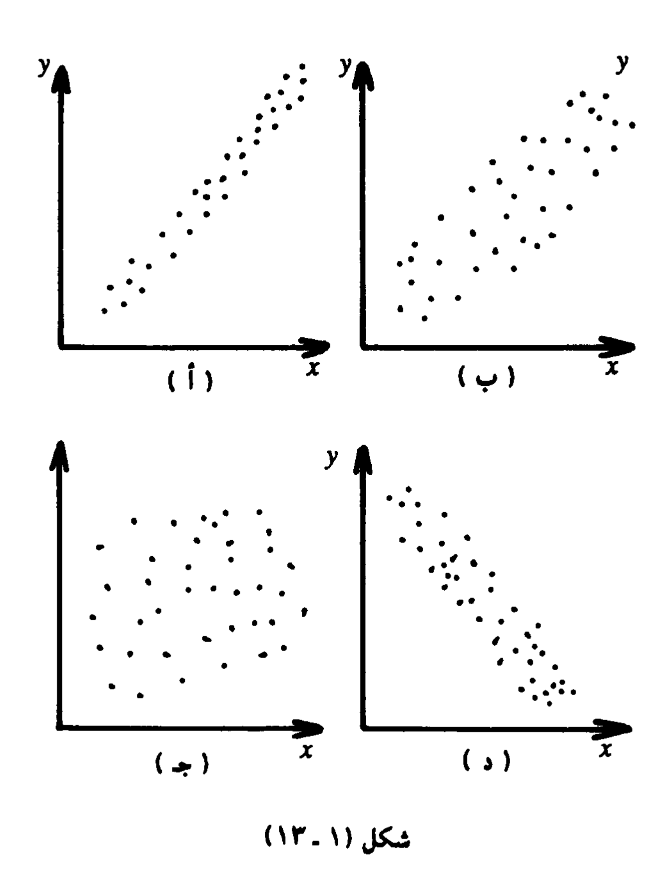
ففي مثل هذه الحالة نقول بوجود ارتباط إيجابي كامل بين المتغيرين x و y، أو بين الظاهرتين اللتين تقيسانها. وإذا وجدنا عند ترتيب القيم أن أصغر قيمة لـx قابلتها أكبر قيمة لـy. والقيمة بعد الصغرى لـx قابلتها القيمة قبل العظمى لـy، . . . ، وأخيرا أكبر قيمة لـx قابلتها أصغر قيمة لـy. فعندئذ نقول بوجود ارتباط سلبي كامل بين المتغيرين x، y؛ أو بين الظاهرتين اللتين تقيسانها. وبين هاتين الحالتين المتطرفتين يمكن أن نتصور ترتيبات تمثل درجات مختلفة من الارتباط في الإتجاه الإيجابي أو في الإتجاه السلبي. ولو أننا سجلنا قيم y على n ورقة صغيرة، وطويناها ثم خلطناها جيدا في جعبة صغيرة، وسحبنا عشوائيا ورقة منها ثم سجلنا القيمة المذكورة فيها أمام أصغر قيمة لـx، وسحبنا ورقة ثانية عشوائيا وسجلنا القيمة المذكورة فيها أمام القيمة بعد الصغرى لـx، وهكذا. . . ، حتى نصل إلى آخر ورقة بقيت في الجعبة فنسجل القيمة

المذكورة فيها أما أكبر قيمة لـx، فيمكن القول، مع مثل هذا الترتيب أو التقابل بين قيم x وقيم y، بعدم وجود أي ارتباط بين الظاهرتين. ويمكن تحري وجود صلة بين المتغيرين برسم أزواج القياسات (x_1, y_1)، (x_2, y_2)، . . . ، (x_n, y_n) بيانيا، حيث قيمة x هي الإحداثي السيني، وقيمة y المقابلة هي الإحداثي الصادي. ونحصل بذلك على n نقطة في مستوى الإحداثيات ونسمي الشكل الحاصل «مخطط الإنتشار». والنظر إلى مخطط الإنتشار يولد نوعا من الانطباع البدهي عن درجة الصلة أو الارتباط القائمة بين المتغيرين.

والشكل (١ ـ ١٣ ـ ا) يمثل حالة ارتباط إيجابي مرتفع، ونلاحظ فيه أن النقاط تحدد اتجاها واضحا وفق خط مستقيم إلى حد كبير. ولو وقعت النقاط بالضبط على استقامة واحدة، لكان الارتباط إيجابيا تاما. والشكل (١ ـ ١٣ ـ ب) يمثل ارتباطا إيجابيا منخفضا إذ يكشف المخطط عن نزعة تأخذ، إلى حد ما، شكل الحزمة الخطية. أما الشكل (١ ـ ١٣ ـ جـ) فيمثل حالة عشوائية، ولا تكشف عن أية نزعات أو اتجاهات واضحة، إذ لا يبدو فيها أي نزوع لاقتران قيم عالية لـx بقيم عالية لـy، وقيم منخفضة لـx بقيم منخفضة لـy. أو العكس، أي قيم عالية لـx بقيم منخفضة لـy وقيم منخفضة لـx بقيم عالية لـy . ويمثل الشكل (١ ـ ١٣ ـ د) ارتباطا سلبيا مرتفعا إلى حد ما، وهنا أيضا، لو وقعت النقاط على استقامة واحدة لكان الارتباط سلبيا تاما. ومن الواضح أنه بين الحالتين المتطرفتين، حالة ارتباط سلبي تام وحالة ارتباط إيجابي تام. يوجد ما لا حصر له ولا عد من إمكانات ترتيب النقاط التي تمثل ما لا حصر له ولا عد من درجات الارتباط الممكنة بين المتغيرين.

ولا بد من التمييز بوضوح بين وجود ارتباط مرتفع بين ظاهرتين وبين وجود علاقة سببية بينهما. فوجود ارتباط مرتفع لا يعني بالضرورة أن إحدى الظاهرتين هي سبب للأخرى؛ إذ قد يكون الارتباط المرتفع بينهما نتيجة لتأثر كل منهما بظاهرة ثالثة لم تدخل في الحساب.

فمثلا، من المعروف أن هناك ارتباطا مرتفعا بين ظاهرة الابتلاء بعادة التدخين والإصابة بمرض سرطان الرئة. وهناك أيضا ارتباط مرتفع بين ظاهرة الابتلاء بعادة



شكل (١ ـ ١٣)

التدخين وتلون أو اصفرار الأسنـان. ولو حصل أن أخذنا بيانا إحصـائيا يتضمن درجة تلون الأسنان ونسبة الإصابة بسرطان الرئة، وكـان هذا البيان في غالبيته من أفراد تلونت أسنانهم بفعل التدخين فسنجد ارتباطا مرتفعا بين ظـاهرة تلون الأسنان وظاهرة الإصابة بسرطان الرئة. وهذا لا يعني بالطبع أن اصفرار الأسنـان يؤدي إلى الإصابة بسرطان الرئة أو العكس، وقد لا يـوجد أي ارتباط إحصـائي فعلي بين الظاهرتين، فـالارتباط المرتفع كان نتيجة لوجود عامل ثالث خفي هو عادة التدخين.

وسنستعـرض الآن إمكـانيـة إيجاد معيـار كمي للتعبير عن درجـة الارتبـاط بين متغيرين نسميه معامل الارتباط.

(١ ـ ١٣ ـ ٢) معامل بيرسون للارتباط

هناك أكثر من صيغة للتعبير عن معامل الارتباط بين متغيرين x و y ؛ ولكنها تعرَّف جميعها لتأخذ قيمها بين 1- تعبيرا عن ارتباط سلبي تام، (وعندئذ تقع جميع النقاط (x, y) على خط مستقيم تتناقص معه قيم y عندما تزداد قيم x، وتتزايد y عندما يتناقص x) وبين 1+ تعبيرا عن ارتباط إيجابي تام. (وعندئذ تقع جميع النقاط (x, y) على خط مستقيم يزداد وفقا له أحد المتغيرين مع زيادة الآخر ويتناقص مع تناقصه) أما القيمة صفر فتعني عدم وجود أي ارتباط أو نزعة أثر أو تأثر بين قيم أحد المتغيرين وقيم المتغير الآخر. ومقياس الارتباط الأكثر استخداما هو معامل بيرسون، ونرمز له عادة بـR.

تعريف معامل بيرسون

لتكن $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ جملة من n من أزواج القياسات. ولنفرض أن $\bar{x}$ و S_x هما متوسط قيم المتغير x وانحرافها المعياري، وأن $\bar{y}$ و S_y هما متوسط قيم المتغير y وانحرافها المعياري. نعرف معامل بيرسون للارتباط بين قيم المتغير x وقيم المتغير y بأنه:

$$R = \frac{1}{n-1} \sum_{i=1}^{n} Z_i \, Z'_i$$

حيث

$$Z'_i = \frac{y_i - \bar{y}}{S_y} \qquad Z_i = \frac{x_i - \bar{x}}{S_x}$$

وقد رأينا في ختام الفقرة السابقة أن معايرة جملة من القياسات تجعل متوسطها صفرا، وتباينها الواحد، وأن مجموع مربعات القيم بعد معايرتها يساوي عدد القياسات في الجملة مطروحا منه الواحد. وهكذا يمكننا كتابة:

$$\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} Z_i'^2 = n-1$$

لنأخذ الآن الحالة الخاصة التي يكون فيها $Z_i = Z'_i$ ، فعندئذ تقع النقاط $(Z_1, Z'_1), \ldots, (Z_n, Z'_n)$ على خط مستقيم هو منصف الربع الأول، ويكون الارتباط في هذه الحالة إيجابيا وتاما. لنحسب R فنجد:

$$R=\frac{1}{n-1}\sum_{i=1}^{n}Z_i\,Z'_i=\frac{1}{n-1}\sum_{i=1}^{n}Z_i^2=\frac{n-1}{n-1}=1$$

وإذا أخذنا الحالة الخاصة المتطرفة المقابلة حيث Z_i و Z'_i متساويان في القيمة المطلقة ومختلفان في الإشارة، فعندئذ تقع النقاط $(Z_1, Z'_1),\ldots,(Z_n, Z'_n)$ على خط مستقيم هو منصف الربع الثاني، ويكون الارتباط في هذه الحالة سلبيا وتاما، أما قيمة R فهي 1-، ذلك لأن:

$$R=\frac{1}{n-1}\sum_{i=1}^{n}Z_i\,Z'_i=\frac{1}{n-1}\sum_{i=1}^{n}Z_i(-Z_i)=-\frac{1}{n-1}\sum_{i=1}^{n}Z_i^2=-\frac{(n-1)}{n-1}=-1$$

ويمكن البرهان، بصورة عامة، أن معامل بيرسون للارتباط يأخذ قيما بين 1 - و 1+ . ويكون 1+ في حالة ارتباط إيجابي تام و 1 - في حالة ارتباط سلبي تام.

مثال (١ ـ ٢٧)

لتكن أزواج القياسات التالية:

x	1	2	3	4	5
y	11	13	15	17	19

احسب معامل بيرسون للارتباط R .

الحل

ننظم الجدول التالي:

جدول (١ ـ ٢٢)

x_i	y_i	$Z_i = \dfrac{x_i - \bar{x}}{S_x}$	$Z'_i = \dfrac{y_i - \bar{y}}{S_y}$	$Z_i\ Z'_i$
1	11	-1.2649	-1.2649	1.60
2	13	-0.6325	-0.6325	0.40
3	0	0	0	0
4	17	0.6325	0.6326	0.40
5	19	1.2649	1.2649	1.60

حيث $\bar{x} = 3$, $S_x = 1.5811$, $\bar{y} = 15$, $S_y = 3.1623$

$$\sum_{i=1}^{5} Z_i\ Z'_i = n-1 = 4$$

$$R = \frac{1}{4}\sum_{i=1}^{5} Z_i\ Z'_i = \frac{4}{4} = 1$$

والجدير بالذكر أن $y = 2x + 9$ وأن النقاط الخمس:
(1, 11), (2, 13) (3, 15), (4, 17), (5, 19) تقع على استقامة واحدة. ونلاحظ أن $Z_i = Z'_i$

(١ ـ ١٣ ـ ٣) حساب معامل الارتباط R

عند حساب معامل الارتباط يشكل رد القياسات إلى شكلها المعياري جهدا حسابيا مطولا لا مسوغ له. ويمكن تطوير الصيغة المعطاة في تعريف معامل بيرسون بعمليات تعويض بسيطة بحيث تأخذ أشكالا مختلفة.

١ ـ بالتعويض عن Z_i ، Z'_i نجد:

$$R = \frac{1}{n-1}\sum_{i=1}^{n} Z_i\ Z'_i = \frac{1}{n-1}\sum_{i=1}^{n} \frac{x_i - \bar{x}}{S_x}\ \frac{y_i - \bar{y}}{S_y}$$

$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1) S_x S_y}$$

$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

وأخيرا

$$R = \frac{\sum_{i=1}^{n} X_i Y_i}{\sqrt{\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i^2}}$$

حيث $X_i = x_i - \bar{x}$ ، $Y_i = y_i - \bar{y}$. ويمكن استخدام العلاقة الأخيرة في الحسابات .

مثال (١ ـ ٢٨)

لدينا أزواج القياسات التالية :

x	1	7	2	3	4	12	11	5	10	5
y	2	5	6	4	1	5	8	2	6	1

احسب معامل بيرسون للارتباط بين x و y.

الحل

ننظم الجدول التالي :

جدول (١ ـ ٢٣). حساب معامل الارتباط باستخدام الانحرافات عن المتوسط

	x_i	y_i	X_i	Y_i	X_i^2	Y_i^2	$X_i Y_i$
	5	1	-1	-3	1	9	3
	10	6	4	2	16	4	8
	5	2	-1	-2	1	4	2
	11	8	5	4	25	16	20
$\bar{x}=6$	12	5	6	1	36	1	6
$\bar{y}=4$	4	1	-2	-3	4	9	6
	3	4	-3	0	9	0	0
	2	6	-4	2	16	4	-8
	7	5	1	1	1	1	1
	1	2	-5	-2	25	4	10
المجموع	60	40	0	0	134	52	84

ولدينا بالتعريف:

$$R=\frac{\sum_{i=1}^{10} X_i \; X_i}{\sum_{i=1}^{10} X_i^2 \sum_{i=1}^{10} Y_i^2}$$

وبالتعويض من السطر الأخير في الجدول (١ ـ ٢٣) نجد:

$$R=\frac{84}{\sqrt{134\times 52}}=0.58$$

٢* ـ ومن المفضل، في الغالب، استخدام صيغة حسابية أخرى تعتمد على القياسات y_i, x_i نفسها. وفي الحقيقة، نجد بسهولة أن:

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{n}(x_i\;y_i-\bar{x}\;\;y_i-x_i\;\bar{y}+\bar{x}\;\;\bar{y})$$

* التفاصيل للقراءة فقط.

$$= \sum_{i=1}^{n} x_i\ y_i - \bar{x} \sum_{i=1}^{n} y_i - \bar{y} \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \bar{x}\ \bar{y}$$

$$= \sum_{i=1}^{n} x_i\ y_i - n\,\bar{x}\ \bar{y} - n\,\bar{y}\ \bar{x} + n\,\bar{x}\ \bar{y}$$

$$= \sum_{i=1}^{n} x_i\ y_i - n\,\bar{x}\ \bar{y}$$

$$= \sum_{i=1}^{n} x_i\ y_i - n \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n^2}$$

$$= \frac{1}{n}\left[n \sum_{i=1}^{n} x_i\ y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)\right]$$

ونعلم أنه يمكن كتابة (انظر الفقرة ١ ـ ٨ ـ ٨):

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{1}^{n} x_i^2 - \frac{\left(\sum_{1}^{n} x_i\right)^2}{n} = \frac{1}{n}\left[n \sum_{1}^{n} x_i^2 - \left(\sum_{1}^{n} x_i\right)^2\right]$$

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \frac{1}{n}\left[n \sum_{1}^{n} y_i^2 - \left(\sum_{1}^{n} y_i\right)^2\right]$$

وبالتعويض في الصيغة الحسابية السابقة نجد:

$$R = \frac{n \sum_{1}^{n} x_i\ y_i - \left(\sum_{1}^{n} x_i\right)\left(\sum_{1}^{n} y_i\right)}{\sqrt{\left[n \sum_{1}^{n} x_i^2 - \left(\sum_{1}^{n} x_i\right)^2\right]\left[n \sum_{1}^{n} y_i^2 - \left(\sum_{1}^{n} y_i\right)^2\right]}}$$

ومع أن مظهر الصيغة معقد، إلا أن جدول الحسابات الضروري لتطبيقها يتضمن خمسة أعمدة فقط، وهي تعتمد كليا على القياسات نفسها، وأسهل صيغة للتطبيق عند توفر آلة حاسبة.

مثال (١ ـ ٢٩)

احسب معامل الارتباط R لأزواج القياسات المذكورة في المثال (١ ـ ٢٦) مستخدما الصيغة التي تعتمد على القياسات مباشرة.

الحل

ننظم الجدول التالي:

وبالتعويض في الصيغة الحسابية التي يمكن أن نكتبها باختصار كما يلي:

$$R=\frac{n\Sigma\ x\ y-(\Sigma x)(\Sigma y)}{\sqrt{\left[n\Sigma x^2-(\Sigma x)^2\right]\left[n\Sigma y^2-(\Sigma y)^2\right]}}$$

$$R=\frac{10\times 288-60\times 40}{\sqrt{\left(10\times 494-60^2\right)\left(10\times 212-40^2\right)}}$$

$$=\frac{480}{\sqrt{1340\times 520}}=0.58$$

(١ ـ ١٣ ـ ٤) معامل سبيرمان لارتباط الرتب

ذكرنا في المقدمة أنه إذا كان لمتغيرين x، y ترتيبان متوازيان أي إذا اتفق ترتيب قيم x مع ترتيب قيم y المقابلة اتفاقا تاما كنا في حالة ارتباط إيجابي تام وإذا كان لهما ترتيبان متعاكسان تماما (أصغر قيمة لـ x قابلتها أكبر قيمة لـ y، والقيمة بعد الصغرى لـ x قابلتها القيمة قبل العظمى لـ y، وهكذا حتى نصل إلى أكبر قيمة لـ x وفي مقابلها أصغر قيمة لـ y) قلنا إن الارتباط سلبي تام. ومعامل سبيرمان لارتباط الرتب يترجم بأمانة هذه الفكرة.

جدول (١ ـ ٢٤): حساب معامل الارتباط باستخدام القياسات نفسها

	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
	5	1	25	1	5
	10	6	100	36	60
	5	2	25	4	10
	11	8	121	64	88
	12	5	144	25	60
	4	1	16	1	4
	3	4	9	16	12
	2	6	4	36	12
	7	5	49	25	35
	1	2	1	4	2
المجموع	60	40	494	212	288

لنفــرض الآن عينــة من قيم x تتضمــن n قيـاسـا ، فكيف نحـدد رتب هـذه القياسات؟ نكتب في عمـود أول الأرقام المتسلسلة من 1 إلى n ، وفي عمود مجاور نرتب قيم x من الأصغـر إلى الأكبر، وفي عمود ثالث نكتـب أمام كل قيمـة لـ x رتبة تساوي الـرقم المتسلسل المقابـل لها. ولكن إذا تكررت إحـدى قيم x أكثر من مـرة فهل نعطي القيمة نفسها رتبا مختلفة؟ وإذا بدا مثل هذا الأمر غير مقبول ، وهو في الحقيقة كذلك ، فكيف نتصرف؟ والجواب واضح بالبـداهة ، ففي مثل هـذه الحالة نعتبر رتبـة كل تكرار لتلك القيمة مساوية للمتوسط الحسابي للأرقام المتسلسلة المقابلة لها. فلنفرض، مثلا، أن الأرقام المتسلسلة 4 ، 5 ، 6 ، 7 في العمود الأول ، قابلها في العمود الثاني 70 ، 70 ، 70 ، 70 ، فتكون الرتبة التي نعطيها لكل من هذه القياسات الأربعة المتساوية هي :

$$\frac{4+5+6+7}{4} = 5.5$$

مثال (١ ـ ٣٠)

لتكن مجموعة القياسات 4، 7، 8، 3، 12، 8، 4، 21، 35، 21، 15، 18، 17، 28، 17، 21 . والمطلوب ترتيب هذه القياسات وتحديد رتبة كل منها .

جدول (١ ـ ٢٥)

رتبة x	قيم x مرتبة	الرقم المتسلسل
1	3	1
2.5	4	2
2.5	4	3
4	7	4
5.5	8	5
5.5	8	6
7	12	7
8	15	8
9.5	17	9
9.5	17	10
11	18	11
13	21	12
13	21	13
13	21	14
15	28	15
16	35	16

وبالطريقة نفسها نرتب قيم y ، وكل رتبة لقيمة من قيم x يوافقها رتبة لقيمة y المقابلة . لنقارن الآن رتب قيم x برتب قيم y المقابلة لها . فلقد كتبنا رُتب x وفق التسلسل الطبيعي ومن الأصغر إلى الأكبر، فما هو الحال بالنسبة إلى تسلسل رُتب y؟ هل حققت ترتيبا موازيا، أي تسلسلا طبيعيا مطابقا لتسلسل رتب x أم طرأ فساد ما على التسلسل الطبيعي لرتب y؟ وما هي درجة أو مدى فساد التسلسل الطبيعي هذا؟ وسنقيس درجة أو مدى فساد التسلسل بالعدد Σd^2 ، حيث d هي رتبة x مطروحا منها رتبة y المقابلة .

وهكذا يمثل Σd^2 مجموع مربعات الفروق بين رتب x ورتب y المقابلة لها. ومن الواضح أن هـذا المقياس لـدرجة فسـاد التسلسل الطبيعي في رتب y سيكـون صفـرا إذا، وفقط إذا تطابقت رتب x مع رتب y المقابلة لها، وعندئذ نكون في حالة ارتباط إيجابي تام. وعندما يكون تسلسل رتب y الناتج بحيث يبدأ بالأكبر وينتهي بالأصغر، أي عكس التسلسل الطبيعي تماما، فإن Σd^2 سيكون أكبر ما يمكن. وهذه الحالة كما أسلفنا هي حالة ارتباط سلبي تام.

ونحتاج الآن إلى تعريف لمعامل ارتباط يعطي القيمة 1+ في الحالة الأولى، والقيمة 1- في الحالة الثانية، ويأخذ القيمة صفرا في حالة عدم وجود أي ارتباط. والمعامل الذي يواجه كل هذه المتطلبات، وسنرمز له بـ τ تمييزا له عن معامل بيرسون للارتباط، هو:

$$\tau = 1 - \frac{2\Sigma d^2}{\text{أكبر قيمة ممكنة لـ } \Sigma d^2}$$

فعندما يتطابق الترتيبان يكون $\Sigma d^2 = 0$ و $\tau = 1$ ، وعندما يتعاكس الترتيبان تماما يأخذ Σd^2 أكبر قيمة ممكنة له، ويكون:

$$\tau = 1 - \frac{2\,(\text{أكبر قيمة ممكنة لـ } \Sigma d^2)}{\text{أكبر قيمة ممكنة لـ } \Sigma d^2} = 1 - 2 = -1$$

ويمكن برهان أنه في حالة عدم وجود ارتباط يكون

$$2\Sigma d^2 = \text{أكبر قيمة ممكنة لـ } \Sigma d^2$$

أي $\tau = 0$.

وإذا كنا ندرس الارتباط في n من أزواج القياسات، فيمكن البرهان على أن أكبر قيمة ممكنة لـ Σd^2 هي $\frac{n(n^2-1)}{3}$ ، وبالتعـويض في العلاقة السابقـة نصل إلى معامل سبيرمان لارتباط الرتب، وهو:

$$\tau = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

مثال (١ ـ ٣١)

لتكن مجموعة الأزواج من القياسات :

x	4	4	7	7	7	9	16	17	21	25
y	8	16	8	8	16	20	12	15	25	20

احسب معامل سبيرمان لارتباط الرتب τ .

الحل

(١) نرتب قيم x ، ثم نرتب قيم y في جدولين منفصلين وفق الطريقة الموضحة في المثال (١ ـ ٣٠).

رتبة y	قيم y مرتبة	الرقم المتسلسل
2	8	1
2	8	2
2	8	3
4	12	4
5	15	5
6.5	16	6
6.5	16	7
8.5	20	8
8.5	20	9
10	25	10

رتبة x	قيم x مرتبة	الرقم المتسلسل
1.5	4	1
1.5	4	2
4	7	3
4	7	4
4	7	5
6	9	6
7	16	7
8	17	8
9	21	9
10	25	10

(٢) ننظم الآن جدولا يتضمن عموده الأول رتب x ، ويتضمن عموده الثاني رتب y المقابلة لها (التقابل بين قيم x وقيم y مبين في المثال). ويتضمن العمود الثالث الفرق d ، وهو يساوي الفرق بين رتبة x ورتبة y المقابلة لها. ومجموع هذا العمود يساوي الصفر، ويتضمن العمود الرابع مربعات الفروق d ، ومجموع هذا العمود هو Σd^2.

رتبة x	رتبة y	d	d^2
1.5	2	-0.5	0.25
1.5	6.5	-5.0	25.0
4	2	2.0	4.0
4	2	2.0	4.0
4	1.5	2.5	6.25
6	8.5	2.5	6.25
7	4	3.0	9.00
8	5	3.0	9.00
9	10	-1.0	1.00
10	8.5	11.5	2.25
المجموع		0	67.00

(٣) نعوض الآن في العلاقة :

$$\tau = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$

حيث $\Sigma d^2 = 67$ ، $n = 10$ فنجد :

$$\tau = 1 - \frac{6 \times 67}{10(100 - 1)} = 0.594$$

تمارين (١ ـ ٥)

١) احسب معامل سبيرمان للارتباط في البيان الإحصائي التالي بعد أن ترسم مخطط الانتشار.

x	y	x	y	x	y	x	y
22	18	35	47	19	37	8	18

تابع :

x	y	x	y	x	y	x	y
15	16	46	22	36	42	1	3
9	31	16	25	25	20	9	7
7	8	7	36	10	12	18	28
4	2	6	27	11	17	46	21
45	36	46	45	5	6	9	25
19	12	11	18	26	45		
26	16	27	18	19	30		

٢) لدى مدرس قناعة بأن قائمة من أسئلة «الخطأ والصواب» ستعطيه من المعلومات عن كفاءة الطلاب في مادته، مثل ما تعطيه مجموعة من الأسئلة تتضمن تمارين ومناقشة. ولكي يثبت وجهة نظره، أعد للطلاب امتحانا يتضمن 25 سؤالا من نوع «الخطأ والصواب»، وما تبقى من الامتحان كان تمارين وأسئلة مناقشة. وقسم العلامة التامة، وهي 200، إلى 50 للقسم الأول (خ، ص)، و 150 للقسم الثاني. وفيما يلي درجات طلابه الثلاثين في كل من القسمين، هل تجد معامل ارتباط مرتفع بين المجموعتين من الدرجات؟ وماذا تستنتج؟ ارسم مخطط الانتشار.

(خ، ص)	تمارين	(خ، ص)	تمارين	(خ، ص)	تمارين	(خ، ص)	تمارين
24	150	21	118	23	125	22	135
23	170	19	110	12	102	14	78
24	141	21	129	15	94	15	105
13	84	25	145	16	91	25	141
19	123	16	124	20	127	19	105
17	100	19	108	21	120	17	110
14	92	18	112	16	105		
18	105	16	98	25	149		

٣) فيما يلي قياس الحذاء x، والوزن بالباوند y، لكل من عشرة طلاب جامعيين :

x قياس الحذاء	9.5	9.5	10.5	10.5	11	8.5	8.5	9.5	10	9
y الوزن	140	155	153	150	180	160	155	145	163	150

أ ـ احسب معامل بيرسون للارتباط ،

ب ـ احسب معامل سبيرمان لارتباط الرتب .

٤) فيما يلي سبعة أزواج من القياسات :

x	10	20	30	40	50	60	70
y	-4	-3	-2	0	3	6	7

ارسم مخطط الانتشار ثم احسب معامل الارتباط .

٥) فيما يلي طول الأم بالبوصة ، x ، وطول ابنتها بالبوصة ، y :

x طول الأم	67	64	62	65	69	63	65	66
y طول الابنة	70	69	65	68	66	60	64	66

ارسم مخطط الانتشار واحسب معامل الارتباط بطريقتي بيرسون وسبيرمان .

٦) سجلنا لعشرة عمال طباعة كلا من معدل إنتاجه في الساعة من الوحدات الجيدة ، x؛ ومعدل إنتاجه في الساعة من الوحدات المعيبة ، y ، فوجدنا ما يلي :

x	94	98	106	114	107	93	98	88	103	95
y	4	5	6	7	6	5	6	4	7	5

احسب معامل الارتباط بين x و y.

٧) في معرض فني يتضمن ثلاثين لوحة رتب محكمان اللوحات حسبما يراه عن درجة نجاحها وأعطى كل منهما الرتبة 1 لأفضل لوحة ، و 2 لتلك التي تليها في الأفضلية حسب رأيه ، وهكذا حتى وصلا إلى 30 لأردأ لوحة كل في رأيه . وفيما يلي الرتب التي

أعطاهـا المحكمان لكل من اللـوحات الثلاثين. احسـب معامل سبيرمـان لارتباط الرتب. ماذا تستنتج؟

رتبة المحكم الأول	1	2	3	4	5	6	7	8	9	10	11
رتبة المحكم الثاني	2	4	3	1	5	7	10	17	8	9	14
رتبة المحكم الأول	12	13	14	15	16	17	18	19	20	21	22
رتبة المحكم الثاني	6	15	11	13	12	18	19	21	16	23	30
رتبة المحكم الأول	23	24	25	26	27	28	29	30			
رتبة المحكم الثاني	29	20	22	25	24	28	26	27			

٨) فيما يلي درجة مادة الـرياضيات x ودرجة مـادة العلوم y لكل من عشرة طـلاب في المرحلة الثانوية:

x	90	95	70	70	65	65	65	40	55	60
y	97	97	85	65	70	70	60	55	40	70

أ ـ ارسم مخطط الانتشار.

ب ـ احسب معامل بيرسون للارتباط بين x و y.

جـ ـ احسب معامل سبيرمان لارتباط الرتب بين x و y.

٩)* فيما يلي تطـور انتــاج القمح x في المملكـة بآلاف الأطنـان وتطـور مجموع القـروض الزراعية الممنوحة y، بملايين الريالات، وذلك بين عامى ١٣٩١هـ و ١٤٠٣هـ.

إنتاج القمح x	42	39	64	153	132	93	125	120	150
مجموع القروض الزراعية y	16.6	16.6	19.6	36.3	145.5	269.4	489.9	585.6	709.1
إنتاج القمح x	142	187	412	741					
مجموع القروض الزراعية y	1128.6	2530.8	2932.9	4166..0					

احسب معامل الارتباط بين x و y.

١٠) يعطي البيان التالي معدلات ما قبل الحرب لامدادات الطعام الصافية x، ومعدلات وفيات الأطفال y في عدد مختار من الدول:

* مأخوذة من منجزات خطط التنمية الصادر عن وزارة التخطيط في المملكة. ص ٢٠٩ وص ٢١٣.

y	x	البلاد	y	x	البلاد	معدل وفيات الأطفال لكل 1000 (y)	عدد الحريرات اليومية للشخص الواحد x	البلاد
32.2	3260	نيوزيلاند	64.2	3420	الدانمرك	98.8	2730	الأرجنتين
40.5	3160	النرويج	162.9	2450	مصر	39.1	3300	استراليا
37.4	3010	هولندا	66.1	2880	فرنسا	87.4	2990	النمسا
139.4	2710	بولونيا	63.3	2960	ألمانيا	83.1	3000	بلجيكا
43.3	3210	السويد	113.4	2600	اليونان	202.1	2080	بورما
45.3	3110	سويسرا	42.4	3160	آيسلند	67.4	3070	كندا
55.3	3100	المملكة المتحدة	161.6	1970	الهند	182.8	1920	سيلان
53.2	3150	الولايات المتحدة	69.6	3390	إيرلندا	240.8	2240	شيلي
94.1	2380	أورغواي	102.7	2510	إيطاليا	155.6	1860	كولومبيا
			60.6	2180	اليابان	116.8	2610	كوريا

ارسم مخطط الانتشار واحسب معامل الارتباط بين عدد الحريرات اليومية للشخص الواحد (x)، وبين y معدل وفيات الأطفال لكل 1000 .

١١) في التمرين ١٨ من مجموعة التمارين (١ ـ ١). معتبرا عدد الأسرة x وعدد الأطباء y . ارسم مخطط الانتشار. واحسب معامل الارتباط بين x و y .

١٢) يتضمن البيان التالي معدل استهلاك الكحول السنوي باللتر للشخص الواحد ممن تزيد أعمارهم عن الرابعة عشرة، x ، ومعدل الوفاة لكل مائة ألف من السكان بمرض تشمع الكبد أو الإدمان، y ، وذلك في مختارات من الدول. ارسم مخطط الانتشار لإيضاح وجود رابطة بين المتغيرين x و y ، ثم احسب معامل الارتباط بينهما.

البلاد	معدل استهلاك الكحول السنوي باللیتر (x)	معدل الوفاة لكل 10^5 من السكان بسبب تشمع الكبد أو الإدمان (y)
فرنسا	24.7	46.1
إيطاليا	15.2	23.6
ألمانيا الغربية	12.3	23.7
استراليا	10.9	7.0
بلجيكا	10.8	12.3
الولايات المتحدة	9.9	14.2
كندا	8.3	7.4
إنكلترة وويلز	7.2	3.0
السويد	6.6	7.2
اليابان	5.8	10.6
هولندا	5.7	3.7
إيرلندا	5.6	3.4
النرويج	4.3	4.3
فنلندا	3.9	3.6

الفصل الثاني

الاحتمــال

(٢ ـ ١) التجارب العشوائية

نـواجـه في معظم ميـادين النشـاط العلمي وفي الحيـاة العمليـة اليـوميـة تجارب ومشاهدات وظـواهر يمكن أن تتكرر عددا كبيرا مـن المرات تحت ظروف متشابهة. وفي كل مـرة نهتم بنتائج هـذه التجارب والمشاهـدات التي يمكن أن تكون كميـة، فنسجل نتيجة كل مشاهدة على شكل عدد. أو قد تأخذ شكلا كيفيا فنسجل صفة معينة كأن نلاحظ مثلا لونا أو نسجل وقوع أو عدم وقوع حادثة أو ظاهرة بعينها متصلة بكل تجربة من التجـارب التي نتابعهـا. ويمكننا، بصـورة عـامة، تعـريف التجربـة على الشكل التالي.

تعريف التجربة

التجربة هي كل عملية تودي إلى ملاحظة «مشاهدة» أو قياس.

ونذكر على سبيل المثال:

(١) عند تكرار رمي حجر نرد عادي نحصل في كل مرة على أحد الأوجه:

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

ويمكن أن نصطلح على تسجيل العـدد 1 نتيجة للتجربة إذا ظهـر الوجه الذي نقشت عليه نقطـة واحدة وتسجيـل العدد 2 إذا ظهـر الوجـه الـذي نقشت عليه نقطتـان، وهكذا. وسنحصل في كل مرة نرمي فيها الحجر على أحد الأعداد

1,2,3,4,5,6,

(٢) عند قياس طول ووزن مجموعة من الأشخاص لكل منهم العمر والجنس نفسيهما فإننا نعبر عن كل ملاحظة بزوج من الأعداد (x,y) فترمز x لقياس الطول و y لقياس الوزن.

(٣) إذا أخذنا عينة من الانتاج اليومي لمصنع من الفولاذ وقسنا في كل قطعة؛ القساوة، المقاومة ، نسبة الفحم فستتألف كل ملاحظة من ثلاثة أعداد.

(٤) إذا تابعنا بشكل دوري منتظم سعر سلعتين معاشيتين، الحليب والبيض، مثلا، فسنعبر عن كل ملاحظة في زوج من الأعداد.

(٥) إذا كنا نتابع جنس كل طفل يولد في منطقة معينة فإننا سنحصل على نتيجة وصفية: ذكر أو أنثى، ويمكن أن نصطلح على التعبير عن هاتين النتيجتين الممكنتين بالرقم 1 أو الرقم 0 ونسجل 1 إذا كان المولود ذكرا و0 إذا كان المولود أنثى.

ونلاحظ في مثل هذه التجارب أن الملاحظات التي نحصل عليها من تكرار للتجربة إلى آخر تعاني تذبذبا عشوائيا لا يخضع لأ ي صيغ أو قوانين معروفة. وبصرف النظر عن العناية القصوى التي نبذلها في كل حالة للتحكم بظروف التجربة ومحاولة إخضاعها لإرادة المجرب، فإن النتائج المتعاقبة ستتغير بشكل غير منتظم من ملاحظة لأخرى، وبصورة تحجب قدرتنا على التنبؤ بالنتيجة سلفا. ونقول في مثل هذه الحالات إننا نقوم بسلسلة من التجارب العشوائية.

وعلى الوجه الآخر، قد نكون في بعض الحالات على درجة كافية من المعرفة الدقيقة بالقوانين التي تتحكم بالظاهرة المدروسة، تبرر لنا التنبؤ الدقيق سلفا بما ستكون عليه نتائج تجربتنا. فإذا كانت التجربة ، مثلا، هي ملاحظة عدد مرات الكسوف الشمسي التي يمكن ملاحظتها من مرصد معين في كل عام، فإننا لا نتردد في القيام بالتنبؤ بهذا العدد، اعتمادا على جداول وحسابات فلكية. وإذا كنا في صدد ملاحظة وتسجيل شدة التيار في دائرة كهربائية، فإننا نستخدم القانون الفيزيائي المعروف:

$$\vartheta = mi$$

حيث ϑ فرق الجهد بين قطبي الدائرة مقاسا بالفولط، و m المقاومة مقاسة بالأوم، و i شدة التيار مقاسة بالأمبير. وهو يسمح لنا بوصف ظاهرة فيزيائية وصفا دقيقا، فنقول مثلا إن دائرة كهربائية، فرق الجهد بين قطبيها 150 فولط، ومقاومتها الكلية 50 أوم، ستكون شدة التيار فيها 3 أمبير. ويبرز نوع مشابه في كل حالة تتوفر لنا فيها معرفة القوانين التي تتحكم بالظاهرة التي ندرسها من جهة، وتكون هذه القوانين، من جهة أخرى، على درجة من البساطة بحيث نتمكن من تطبيقها عمليا.

والخاصة المميزة للتجارب العشوائية هي التذبذب غير المنتظم في نتيجة التجربة من تكرار إلى آخر، وبالنسبة إلى تجربة عشوائية يجب أن يكون في مقدورنا تحديد مجموعة كل النتائج التي يمكن أن يسفر عنها تنفيذ التجربة مرة واحدة، إلا أنه لا يمكن التنبؤ سلفا بالنتيجة التي سنحصل عليها من بين تلك المجموعة من النتائج الممكنة.

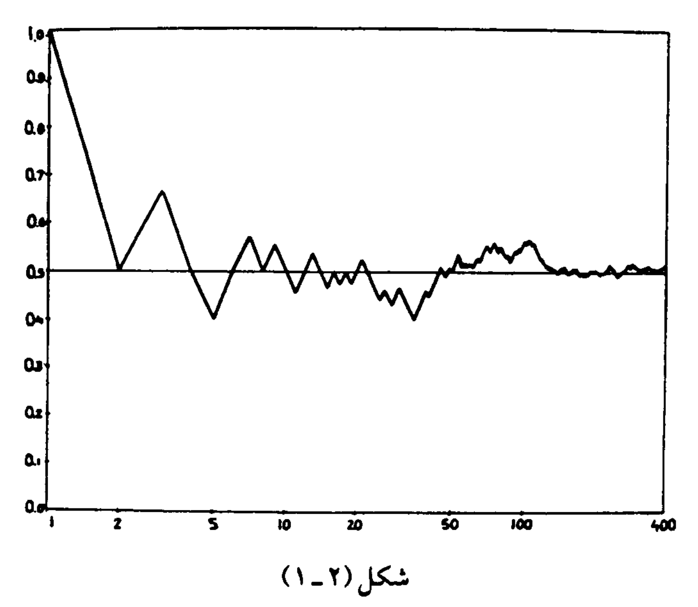
وسنرى الآن أنه في وسط هذا التذبذب غير المنتظم الذي تتصف به التجارب العشوائية، يبدو لنا خيط من الأمل، يتمثل في ظاهرة نزوع نحو الانتظام على المدى البعيد.

(٢ ـ ٢) الانتظام الإحصائي

رأينا أنه لا يمكننا التنبؤ بنتيجة تجربة بمفردها عند القيام بسلسلة من التجارب العشوائية، وأن النتائج المتتابعة لتكرار التجربة تحت الشروط نفسها تخضع لتذبذبات عشوائية غير منتظمة، إلا أننا عندما تحول اهتمامنا من التجارب واحدة فأخرى، إلى مجمل السلسلة من التجارب التي أجريناها ككل، فإن الأمر يختلف كليا، وتبدو لنا ظاهرة مهمة جدا، نعبر عنها على الشكل التالي: بالرغم من السلوك غير المنتظم للنتائج مفردة، فإن معدل هذه النتائج في سلسلة طويلة من التجارب يُظهر انتظاما مدهشا.

ولايضاح الفكرة، نأخذ تجربة قذف قطعة نقود، وسنرمز بـ H لوجه الصورة، وبـ T لوجه الكتابة. إذا كررنا التجربة 20 مرة، مثلا، ورأينا أن وجه الـ T قد ظهر في 12 منها، قلنا إن التكرار النسبي لحادثة ظهور الوجه T هو 12/20

وبصورة عامة، إذا كررنا التجربة N مرة وظهر وجه ال H في n منها فإن التكرار النسبي لظهور وجه ال H هو n/N . ويوضح الشكل (٢ ـ ١) كيف يتغير التكرار النسبي n/N مع قيم متزايدة لعدد التكرارات N . وكما نرى على الشكل يتذبذب التكرار النسبي بشدة من أجل قيم صغيرة لـ N . ولكن هذا التذبذب يصبح أضعف فأضعف مع زيادة N. ويثير هذا الشكل الانطباع بأنه إذا أمكن زيادة العدد N بلا حدود، أي أمكن تكرار التجربة تحت الشروط نفسها بلا تناه، فإن التكرار النسبي سيسعى إلى نهاية قريبة جدا من النصف.



شكل (٢ ـ ١)

والخبرة التجريبية تشير، على وجه العموم، إلى أن التكرارات النسبية تسعى إلى الاستقرار، عادة، بعد سلسلة طويلة من المشاهدات العشوائية المتكررة التي تجري تحت شروط منتظمة. ونظرة فاحصة عن كثب إلى الحالات التي يبدو فيها وكأن مثل هذا النزوع إلى الا ستقرار غير صحيح، ستزيح الستار عن نقص أكيد في انتظام الشروط التي نكرر تحتها التجربة. وهذا يدفعنا إلى القول إنه إذا أمكن الاستمرار في سلسلة لا نهاية لها من التكرارات لتجربة عشوائية E ، مثلا، وسجلنا في كل تكرار وقوع أو عدم وقوع حادثة E ، مثلا، مرتبطة بهذه التجربة، وراقبنا تطور قيمة التكرار النسبي لوقوع الحادثة E ، فسنرى أنه يسعى، بصورة عامة، إلى قيمة مثالية محددة. وبالطبع فإنه لا

يمكننا برهان صحة أو عدم صحة هذه المقولة، طالما أنه لا يمكننا أصلا القيام بسلسلة من التكرارات لا نهاية لها. إلا أن التجارب تؤيد، بصورة عامة، المقولة التالية الأقل دقة، وهي أنه يمكننا أن ننسب إلى كل حادثة E مرتبطة بتجربة عشوائية F ، عددا p ، حتى إذا قمنا بسلسلة طويلة من التكرارات للتجربة يصبح التكرار النسبي لوقوع الحادثة E مساويا تقريبا لـ p . وهذه هي الصيغة النموذجية للانتظام الإحصائي الذي يشكل الأساس التجريبي لنظرية الإحصاء.

(٢ ـ ٣) هدف النظرية الرياضية

عندما نكتشف في مجموعة من الظواهر التي يتطرق إليها النشاط الإنساني، عن طريق الملاحظة والتجربة، دلالات كافية على نوع من الانتظام، فإن هذا يدفعنا إلى بلورة نظرية رياضية لمثل هذه الظواهر تشكل النموذج الرياضي أو القالب الذي يحتوي على الحقائق العملية كافة المستوحاة من معطيات الملاحظة والتجربة.

وعندئذ تكون نقطة البداية هي أن نختار أكثر حقائق هذا الانتظام بساطة وجوهرية ونصوغها، على شكل مبسط من جهة ومجرد ومثالي من جهة أخرى، كموضوعات رياضية تشكل المسلمات أو البديهيات التي نبني عليها نظرية رياضية، ونعتبرها في مستوى الحقائق المسلم بها سلفا، ثم نستنتج انطلاقا من هذه المسلمات موضوعات أخرى لا نحتاج في عملية استخلاصها إلى غير المنطق الرياضي المجرد، ودون أية حاجة إلى العودة إلى معطيات الملاحظة والتجربة. ويشكل مثل هذا البناء الذي نستخدم فيه الاستنتاج المنطقي وحده، والذي يتعاظم يوما بعد يوم من خلال جهود البحث والاستقصاء، ما يسمى بالنظرية الرياضية.

وكل موضوعة صحيحة تماما من وجهة النظر الرياضية طالما استنتجناها بصورة منطقية من المسلمات. إن النقاط والمستقيمات والمستويات الخ. التي ترد في علوم الهندسة البحتة هي تجريدات ذهنية لا وجود لها في الواقع. والنظرية البحتة تنتمي بشكل كامل إلى دائرة الأفكار المجردة، وتعالج أشياء وموضوعات مجردة ومعرفة تماما

بالخواص الممنوحة لها من قبل المسلمات. وعلى سبيل المثال، فالموضوعة الإقليدية بأن مجموع زوايا المثلث يساوي π راديان هي موضوعة صحيحة تماما في صورة مجردة ذهنية للمثلث كما تعرفه الهندسة البحتة. ولكن هذا لا يعني أن مجموع زوايا مثلث واقعي، أو مثلث نرسمه على الورق، يساوي π تماما.

وعلى أية حال، يمكن اختبار قضايا معينة من نظرية رياضية عمليا. إذ يمكن، مثلا، مقارنة الموضوعة المتعلقة بمجموع زوايا مثلث بقياس حقيقي لمجموع زوايا مثلث واقعي، وإذا حققت الاختبارات المتتالية، وإلى درجة كافية ومرضية من الدقة، توافقا بين النظري والواقعي، قلنا إن هناك نوعا من التشابه بين النظرية الرياضية وبناء العالم الواقعي. ونتوقع فوق هذا أن مثل هذا التوافق سيبقى قائما ومستمرا في المستقبل، سواء فيما تم اختباره، أو فيما لم يتعرض بعد لامتحان الواقع. ونسمح لأنفسنا بالسير على هدى مثل هذا التوقع. وتستمد النظرية قيمتها العملية مما يتوفر لنا من أدلة على التوافق الدقيق والدائم بينها وبين حقائق العالم الواقعي.

والحساب الاحتمالي هو النظرية التي تشكل النموذج الرياضي للظواهر التي تتصف بالانتظام الإحصائي. وسنقدم في هذا الفصل طريقة لبناء النظرية الاحتمالية باعتبارها نظرية رياضية، وذلك في حالة بسيطة وممتعة هي في متناول الطالب المبتدىء في دراسة الإحصاء والاحتمال وهي حالة فضاء عينة منته.

(٢ ـ ٤) فضاء العينة و الحادثة

نفترض دائما أننا قادرون على تحديد كل النتائج التي يمكن أن تسفر عنها التجربة العشوائية لو أننا نفذناها مرة واحدة. وسنطلق على مجموعة كل النتائج الممكنة لتجربة مصطلح «فضاء عينة». وسيمثل كل عنصر من هذه المجموعة (أي كل نتيجة ممكنة للتجربة) نقطة في فضاء العينة أو اختصارا «نقطة عينة». ومن البديهي أنه يمكن التعبير عن أي حادثة تتصل بالتجربة بدلالة نقاط العينة (أي بدلالة النتائج الممكنة للتجربة). وسنرمز لفضاء عينة بـ S .

تعريف فضاء العينة

فضاء العينة هو مجموعة كل النتائج الممكنة لتجربة .

تعريف الحادثة

الحادثة هي مجموعة جزئية من فضاء عينة .

ويتضح من هذين التعريفين أننا لن نتحدث أبدا عن الاحتمالات إلا في علاقتها مع فضاء عينة معطى (أي في علاقتها بتجربة عشوائية معينة) . وأن كل ما يمكن أن نسميه «حادثة» في نظرية الاحتمال يجب أن يكون مجموعة جزئية من فضاء عينة . لقد أصبح لكلمة «الحادثة» الآن معنى جديد يضاف إلى المعاني اللغوية التي نعرفها سابقا . فهي الآن مصطلح رياضي شأنها شأن المستقيم والسطح في الهندسة والمتجه والقوة في الميكانيكا والدالة والسلسلة في التحليل والزمرة والحلقة في الجبر إلى آخره . الحادثة ببساطة هي كائن رياضي مقترن على الدوام باحتمال .

ولغايات التوضيح وتيسير الفهم سيكون مفيدا أحيانا رسم مصور بياني يسمى مصور ڤن لفضاء عينة S ، وذلك بتمثيل كل نقطة عينة كنقطة هندسية ثم إحاطتها بخط مغلق .

مثال (٢ ـ ١)

التجربة هي قذف حجر نرد وملاحظة عدد النقاط المنقوشة على الوجه الظاهر .

ولكتابة فضاء العينة نجيب على السؤال التالي :

إذا قذفنا حجر النرد مرة واحدة فماذا يمكن أن تكون النتيجة؟

والجواب واضح فالنتيجة إما أن تكون 1 أو 2 أو 3 أو 4 أو 5 أو 6 . ويكون :

$$S = \{1, 2, 3, 4, 5, 6\}$$

وبعض الحوادث التي يمكن إيرادها هي ، على سبيل المثال لا الحصر،
١ ـ الحصول على عدد زوجي ، ولنرمز لهذه الحادثة بـ A ،
٢ ـ الحصول على عدد أكبر من 4، ولنرمز لهذه الحادثة بـ B ،
٣ ـ ملاحظة العدد 1 ، ولنرمز لهذه الحادثة بـ E_1 ،
٤ ـ ملاحظة العدد 2 ، ولنرمز لهذه الحادثة بـ E_2 ،
٥ ـ ملاحظة العدد 3 ، ولنرمز لهذه الحادثة بـ E_3 ،
٦ ـ ملاحظة العدد 4 ، ولنرمز لهذه الحادثة بـ E_4 ،
٧ ـ ملاحظة العدد 5، ولنرمز لهذه الحادثة بـ E_5 ،
٨ ـ ملاحظة العدد 6 ، ولنرمز لهذه الحادثة بـ E_6 ،

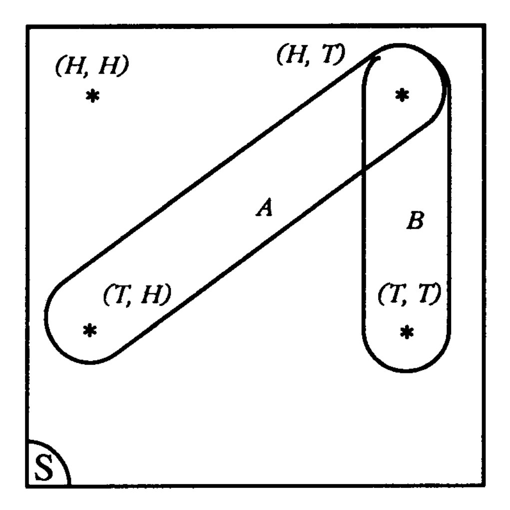
ونلاحظ الفرق بين الحادثتين A و B من جهة والحوادث E_1، E_2، E_3، E_4، E_5، E_6 من جهة أخرى . فستقع الحادثة A إذا وقعت أي من الحوادث E_2، E_4، E_6 ، أي عندما نلاحظ 2 أو 4 أو 6 وهكذا يمكن تفكيك الحادثة A إلى مجموعة من الحوادث الأبسط ، ونقصد E_2 ، E_4 ، E_6 . وكذلك ستقع الحادثة B إذا وقعت أي من الحوادث E_5 أو E_6 ويمكن النظر إليها كمجموعة من الحوادث الأبسط . وفي المقابل نلاحظ أنه من المستحيل تفكيك أي من الحوادث E_1، E_2، E_3، E_4، E_5، E_6 ، أي التعبير عن أي منها بدلالة حوادث أبسط . وهكذا تسمى الحوادث E_1، E_2، E_3، E_4، E_5، E_6 ، حوادث بسيطة (أو حوادث ابتدائية) وحوادث مثل A ، B حوادث مركبة .

وتبدو بوضوح خاصة مهمة من خواص الحوادث البسيطة وهي أن تنفيذ التجربة يؤدي إلى واحدة وواحدة فقط من الحوادث البسيطة ، فعندما نقذف حجر النرد سنحصل حتما على 1 أو 2 أو 3 أو 4 أو 5 أو 6 ، ولا يمكن أن نلاحظ في الوقت نفسه أكثر من واحدة من هذه الحوادث البسيطة .

وبصورة عامة ، كل نقطة عينة بمفردها من فضاء عينة S هي بالطبع مجموعة جزئية من S ، أي حادثة ، ومثل هذه الحوادث سنسميها دائما حوادث بسيطة أو حوادث ابتدائية .

وبما أن $S \subset \phi$ و $S \subseteq S$ فإن تعريف الحادثة ينطبق أيضا على المجموعة الخالية ϕوعلى فضاء العينة S . وتسمى ϕالحادثة المستحيلة و S الحادثة الأكيدة. ومن الواضح أنه يمكن التعبير عن أي حادثة غير مستحيلة (غير الحادثة ϕ) بدلالة حوادث بسيطة.

مثال (٢ ـ ٢)



شكل (٢ ـ ٢). مصوّر ڤن لتجربة قذف قطعة نقود مرتين.

التجربة هي قذف قطعة نقود مرتين متتاليتين وتسجيل النتيجة.

ا) اكتب فضاء العينة

ب) ارسم مصور ڤن

ج) عبّر عن كل من الحوادث التالية بدلالة نقاط العينّة.

A : الحصول على وجه الـ H مرة واحدة.

B : الحصول على وجه الـ T في القذفة الثانية،

C : الحصول على وجه الـ H مرة واحدة على الأقل،

D : الحصول على وجه الـ H مرة واحدة على الأكثر،

E : الحصول على وجه الـ H مرة واحدة على الأقل وعلى وجه الـ T مرتين.

الحــــــل

فضاء العينة هو مجموعة كل النتائج الممكنة عند تنفيذ التجربة مرة واحدة. أي للحصول على فضاء العينة اسأل نفسك السؤال التالي: لو أنني قذفت قطعة نقود مرتين فما هي النتائج التي يمكن أن أحصل عليها؟

ونرمز للنتائج عادة باختصار مستخدمين الرمزين H و T في أزواج مرتبة حيث يرمز الحرف الأول لنتيجة القذفة الأولى والحرف الثاني لنتيجة القذفة الثانية.

والنتائج الممكنة هي:
(H, H) أي وجه الـ H من القذفة الأولى و H من القذفة الثانية،
(T, H) أي وجه الـ T من القذفة الأولى و H من القذفة الثانية،
(H, T) أي وجه الـ H من القذفة الأولى و T من القذفة الثانية،
(T, T) أي وجه الـ T من القذفة الأولى و T من القذفة الثانية،
ويكون فضاء العينة:

$$S = \{(H, H), (T, H), (H, T), (T, T)\}$$

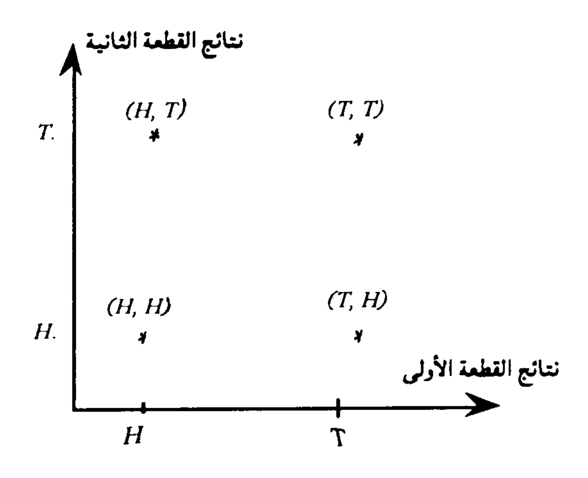
وهذا يكافىء قولنا:

القذفة الأولى يمكن أن تسفر عن H أو T نضعهما في وضع رأسي فوق بعضهما ونضع إشارة استفهام في الموضع الثاني المخصص لنتيجة القذفة الثانية ثم نستعيض عن إشارات الاستفهام مرة بـ H ومرة بـ T لنجد:

$$(H, H), (H, T)$$
$$(T, H), (T, T)$$

أو يمكن تمثيل نتائج القذفة الأولى على محور السينات ونتائج القذفة الثانية على محور الصادات ثم تحديد فضاء العينة المطلوب كبيان لحاصل الجداء الديكارتي للمجموعة (H, T) في نفسها، (انظر الشكل ٢ ـ ٣).

ولكتابة حادثة بدلالة نقاط العينة، أي كمجموعة جزئية من S، نلاحظ أن وصف الحادثة يتضمن شروطا أو مواصفات معينة. ووفقا لهذه الشروط سنجد، بالنسبة إلى كل نقطة عينة، أنها إما أن تحقق هذه الشروط أو المواصفات، وبالتالي تنتمي إلى



شكل (٢ ـ ٣) تمثيل فضاء العينة بيانيا

الحادثة، أو أنها لا تحقق الشروط المطلوبة وبالتالي لا تنتمي إلى الحادثة. وفي الحادثة A نجد أنها تتضمن كل زوج مرتب في S يحوي الرمز H مرة واحدة (لا أكثر ولا أقل). وهكذا نكتب:

$$A = \{(H, T), (T, H)\}$$

أما (H, H) و (T, T) فلا تنتميان إلى A لأنهما لا تحققان شروطها، ولو أننا نفذنا التجربة وحصلنا على نقطة عينة (نتيجة ممكنة) تنتمي إلى A، أي حصلنا على (H,T) أو (T, H)، فسنقول عندئذ إن A قد وقعت. ولو حصلنا على نتيجة أو نقطة عينة لا تنتمي إلى A فسنقول إن الحادثة A لم تقع. وبالطريقة نفسها نجد أن:

$$B = \{(H, T), (T, T)\}$$

$$C = \{(H, H), (T, H), (H, T)\}$$

$$D = \{(T, H), (H, T) (T, T)\}$$

$$E = \{\ \} = \emptyset$$

لاحظ أنه لا توجد أي نقطة عينة محققة لشروط E فهي حادثة غير ممكنة أو مستحيلة.

ونلاحظ أنه لو كانت التجربة قذف قطعة نقود ثلاث مرات فإن الرسم البياني سيحتاج إلى ثلاثة محاور ويصبح تطبيق طريقة الرسم معقدا. ومع أربع قذفات لا تعود

طريقـة الرسم البيـاني مجدية. ولكن الطـريقة المذكورة أولا تبقى صـالحة للتطبيق. ففي تجربة ثلاث قـذفات يكون عدد النتـائج الممكنة $2^3 = 8$. ونحصل عليها بكتـابة النتائج الأربع من أجل قذفتين، وتكرارها مرة مع إضافة H ثم أخرى مع إضافة T. وفي تجربة أربع قذفات نكـرر النتائج الثماني لثلاث قذفات مـرة مع إضافة H ومرة مع إضـافة T لنحصل على النتائج الست عشرة الممكنة في هذه الحالة، وهكذا. . .

مثال (٢ ـ ٣)

التجربة هي قذف حجر نرد مرتين.

ا ـ اكتب فضاء العينة،

ب ـ عبر عن الحوادث التالية بدلالة نقاط العينة.

A: الحصول على مجموع يساوي 7 ،

B: الفرق بين العددين الناتجين يساوي بالقيمة المطلقة 1.

C : الحصول على مجموع يساوي 9 على الأقل،

D : الحصول على 1 في القذفة الأولى،

E : الحصول على جداء يساوي 6 على الأكثر،

F: الحصول على مجموع أقل من 2.

جـ ـ عبر بكلمات عن كل من الحوادث الممثلة بـالمجموعات الجزئيـة التاليـة من نقاط العينة:

$G = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$

$H = \{(1, 1), (2, 1), (1, 2), (3, 1), (2, 2), (1, 3)\}$

$I = \{(5,1), (1, 5), (6, 2), (2, 6)\}$

$J = \{(1,4), (2,4), (3,4), (4,4), (5,4), (6,4)\}$

$K = \{(2, 2), (2, 4), (2, 6), (4,2), (4, 4), (4, 6), (6, 2), (6, 4), (6, 6)\}$

د ـ لو نفـذنا التجربة وحصلنـا على النتيجة $(1, 1)$، حدد وقـوع أو عدم وقـوع كل من الحوادث المذكورة في ب و جـ.

الحـــــل

ا ـ فضاء العينة هو الحاصل الديكارتي للمجموعة {1,2,3,4,5,6} في نفسها، وهو كما في الجدول (٢ـ١).

جدول (٢ ـ ١) . فضاء العينة لقذف حجر نرد مرتين

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

حيث يرمز الزوج المرتب $(x\,,y)$ إلى أن النتيجة كانت x من القذفة الأولى و y من القذفة الثانية. وكان يمكن التعبير عن هذه النتائج الست وثلاثين على الشكل التالي:

$S = \{(x, y):$ x و y عددان صحيحان بين 0 و 7 $\}$

وبدلا مـن الجـدول (٢ ـ ١) كـان يمكن رسم بيـان الحاصل الديكـارتي واعتماده تمثيلا لفضـاء العينـة. ويتم ذلك كما في الشكل (٢ ــ ٤) حيث تتمثل كل زوج مـرتب (كل نقطة عينة) مـن الأزواج الستة وثلاثين المذكورة في الجدول (٢ ــ ١) بنقطة في المستوى، إحداثيهـا السيني هـو العدد الأول من الـزوج المرتب، وإحداثيهـا الصادي هـو العدد الثاني.

ب ـ

$A = \{(6,1), (5,2), (4,3), (3,4), (2,5), (1,6)\}$

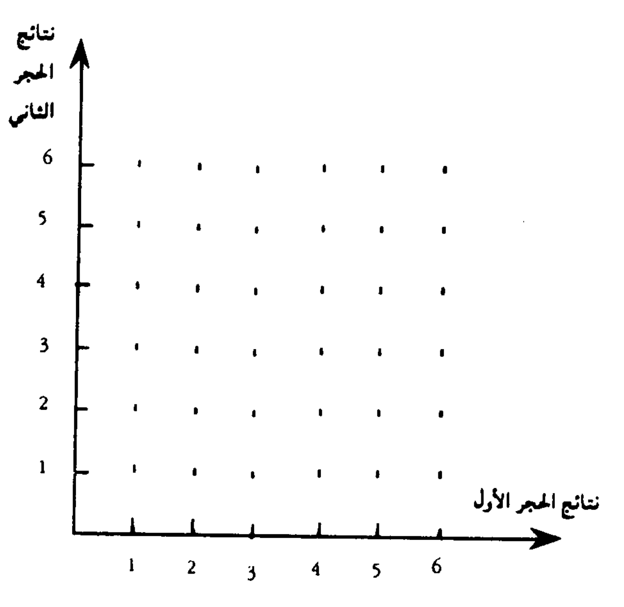
$B = \{(2,1), (1,2), (3,2), (2,3), (4,3), (3,4), (5,4), (4,5), (6,5), (5,6)$

$C = \{(6,3), (5,4), (4,5), (3,6), (6,4), (5,5), (4,6), (6,5), (5,6), (6,6)\}$

$D = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)\}$

$E = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (3,1), (3,2),$
$(4,1), (5,1), (6,1)\}$

$F = \{\ \} = \phi$



شكل (٢ ـ ٤) تمثيل بياني لفضاء العينة في تجربة قذف حجر نرد مرتين

جـ ـ

G : الحصول على العدد نفسه في القذفتين،

H : الحصول على مجموع يساوي 4 على الأكثر،

I : الفرق بين العددين الناتجين يساوي بالقيمة المطلقة 4،

J : الحصول على 4 في القذفة الثانية،

K : الحصول على عددين زوجيين.

د ـ تقع الحادثة أو لا تقع وفقا لما إذا كانت نقطة العينة (1,1) تنتمي أو لا تنتمي إلى الحادثة، أو ما إذا كانت النتيجة «واحد من القذفة الأولى وواحد من القذفة الثانية» تحقق شروط ومواصفات الحادثة. وهكذا نجد أن:

A لم تقع لأن المجموع الناتج (وهو 2) لا يساوي 7،

B لم تقع لأن الفرق بين العددين الناتجين (وهو صفر) لا يساوي 1 بالقيمة المطلقة،

C لم تقع لأن المجموع أقل من 9،

D وقعت لأن القذفة الأولى أنتجت 1 ،
E وقعت لأن جداء العددين الناتجين لا يزيد على 6،
F لم تقع بالطبع لأنها مستحيلة،
G وقعت لأن $(1,1) \in G$،
H وقعت لأن $(1,1) \in H$،
I لم تقع لأن $(1,1) \notin I$،
J لم تقع لأن $(1,1) \notin J$،
K لم تقع لأن $(1,1) \notin K$،

مثال (٢ ـ ٤)

في عملية استطلاع لنسبة المؤيدين لقضية معينة قوبل شخصان، إذا كانت إجابة كل منهما هي إما «مع» وسنرمز لها بـ 1 أو «حيادي» وسنرمز لها بـ 0 ، أو «ضد» وسنرمز لها بـ 1-.

ا ـ اكتب فضاء العينة لهذه التجربة وارسمه بيانيا متخذا المحور الأفقي لإجابة الشخص الذي قوبل أولا، والمحور الرأسي لإجابة الشخص الآخر.

ب ـ عبر كلاميا عن كل من الحوادث الممثلة بالمجموعات التالية من نقاط العينة :

$A = \{(1, 1), (1, 0), (1, -1)\}$

$B = \{-1, -1), (0, 0), (1, 1)\}$

$C = \{(-1, 0), (1, 0), (0, -1), (0, 1), (0, 0)\}$

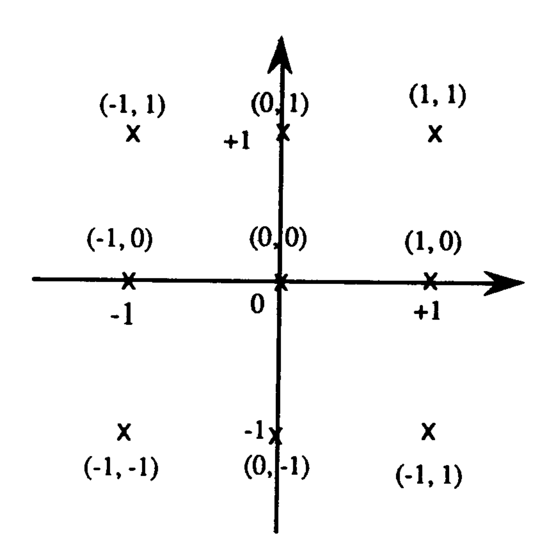
جـ ـ عبر عن الحوادث التالية بدلالة نقاط العينة :
U : الشخص الثاني ضد القضية،
T : واحد منهما على الأقل ضد القضية،
V : أحدهما مع القضية والآخر ضدها .

الحـــل

ا ـ فضاء العينة هو الجداء الديكارتي للمجموعة $\{-1,0,1\}$ في نفسها . أي

$S = \{(-1, -1), (-1, 0), (-1, 1), (0, -1), (0, 0), (0, 1), (1, -1), (1, 0), (1, -1)\}$

والرسم كما في الشكل المقابل



ب ـ

A : حادثة أن الشخص الأول مع القضية ،

B : حادثة أن للشخصين الموقف نفسه ،

C : حادثة أن واحدا منهما على الأقل حيادي .

جـ ـ

$U = \{(-1, -1), (0, -1), (1, -1)\}$

$T = \{(-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1)$

$V = \{(-1, 1), (1, -1)\}$

مثال (٢ ـ ٥)

التجربة هي قذف قطعة نقود حتى يظهر وجه الـ H لأول مرة . اكتب فضاء العينة .

الحـــل

يتضمن فضاء العينة عددا غير محدود من النقاط نلاحظ بوضوح أنها كما يلي :

$H, TH, TTH, TTTH, TTTTH, \ldots\ldots$

فقد لا نحتاج إلا إلى قذفة واحدة حتى يظهر وجه ال H وتنتهي التجربة، وقد نحتاج إلى قذفتين حتى يظهر وجه الـ H للمرة الأولى أو إلى ثلاث قذفات، أو إلى أربع، الخ . . .

مثال (٢ ـ ٦)

التجربة هي اختيار أسرة بصورة عشوائية وتسجيل عمر الزوج x ثم عمر الزوجة y. اكتب فضاء العينة وعبر عن حادثة «الزوج أكبر سنا من الزوجة».

الحـــل

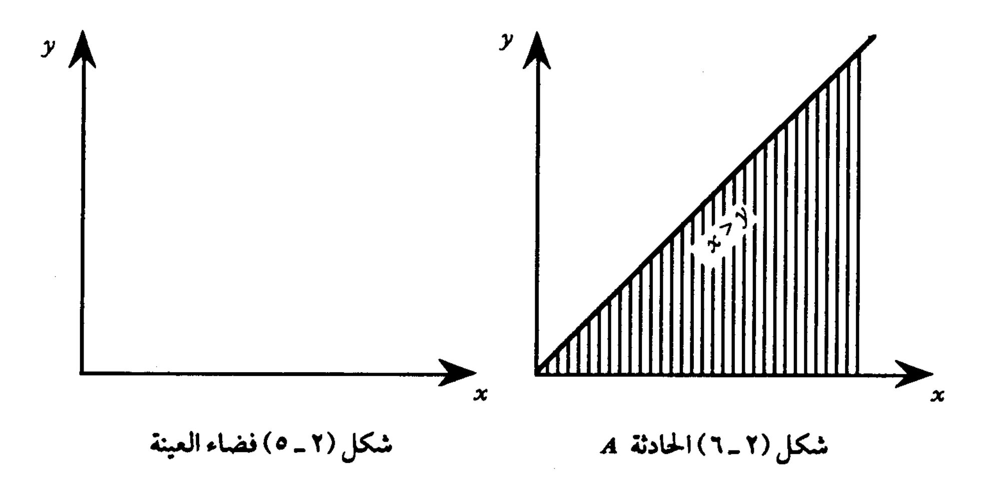
العمر هو قياس زمني يمثل نقطة على محور الزمن ويمكن وصفه بصورة عامة أنه عدد حقيقي موجب، أي ينتمي إلى R^+ حيث يرمز R^+ لمجموعة الأعداد الحقيقية الموجبة. ويمكن التعبير عن فضاء العينة بأنه مجموعة كل الأزواج المرتبة (x, y) حيث x و y عددان حقيقيان موجبان [انظر شكل (٢ ـ ٥)].

$$S = \{(x, y): x, y \in R^+\}$$

وبيانيا نجد أن S هو مجموعة نقاط الربع الأول من مستوى الإحداثيات. وإذا رمزنا لحادثة «الزوج أكبر سنا من الزوجة» بـ A فتكون:

$$A = \{(x, y): x > y;\ x, y \in R^+\}$$

وبيانيا تتضمن الحادثة A كافة نقاط الربع الأول من مستوى الإحداثيات الواقعة تحت منصّف الربع الأول [انظر الشكل (٢ ـ ٦)].



شكل (٢ ـ ٥) فضاء العينة

شكل (٢ ـ ٦) الحادثة A

ومن الواضح أن فضاء العينة S كما حددناه في المثال (٢ ـ ٦) يتضمن من النقاط أكثر بكثير مما يمكن أن نـواجهه بالفعل في الواقع العملي. إذ يمتـد عمر كل من الزوج والزوجة بين عـددين مألوفين ولا يمتد عمليـا بين الصفر واللانهاية، وقـد يبدو في الأمر بعض الغرابة إلا أنها في الواقع غـرابة مقبولة ولابد منها لأنها تتفـادى ، من جهة، ما هو أشـد غرابـة، لا بل معضلة تفـوق قدرتنـا. ولا تقدم، من جهـة أخرى، أذى لبنـاء النظـرية الاحتماليـة بل تجعل هـذا البنـاء أكثر يسرا وسهـولـة، ولإيضاح المعضلـة التي نواجهها عند محاولة تحديد حـد أدنى وحد أعلى لعمر الزوج، مثلا، يكفي أن نتساءل: هل يمكن الادعاء أن عمر الزوج يمكن أن يكون 150 عاما ، مثلا، ولكنه لا يمكن أن يكون 150 وثانيـة واحدة؟ وهل يمكـن الادعاء بأن عمـر الزوجة يمكـن أن يكون عشر سنوات إلا أنه لا يمكن أن يكون عشر سنوات ناقصا ثانيتين؟

وبصورة عامة نقول إنه عند تحديد فضاء عينة لا ضير في أن يتضمن فضاء العينة من النقاط أكثر مما ينبغي عمليا. إلا أنـه لا يجوز أبدا أن يتضمن أقل مما ينبغي عمليا. أي لا يجوز أن نغفل ذكر أو شمـول أي نتيجة ممكنة عمليـا. وعندما نصف العمـر بأنه عـدد حقيقي مـوجب نكون مطمئنين إلى أننـا لم نغفل أي نتيجـة ممكنـة إذ لا يمكن أن يكون العمر سـالبا. وفي الوقت نفسه نتفـادى تحديد حد أدنى وحـد أعلى للعمر، فالله وحده سبحانه وتعالى يعلم، ولا يحيط مخلوق بشيء من علمه إلا بما شاء.

(٢ ـ ٥) جبر الحوادث

عرفنا الحادثة كمجموعة جزئية من فضـاء عينة، أي مجموعة عناصرها نقاط عينة أو نتائج ممكنة لتجربـة عشوائية. وكل ما يعرفه الطـالب عن عمليات الاتحاد والتقاطع والفرق مطبقة على المجموعات، وعن الخواص المختلفة لهذه العمليات، ينسحب تماما على الحوادث بعد أن نضع كلمة «حادثة» بدلا من كلمة مجموعة. وسنستعرض في هذه الفقرة، على سبيل التذكير، هذه العمليات بلغة الحوادث ونقاط العينة.

(٢ ـ ٥ ـ ١) اتحاد حادثتين

اتحاد حادثتين A، B هو حادثة تتضمن كافة نقاط العينة التي تنتمي إلى A أو إلى B (أو إليهما معا). ونرمز له بـ $A \cup B$.

ونلاحظ في هذا التعريف أن شرط انتماء نقطة عينة إلى الاتحاد $A \cup B$ هو أن تنتمي هذه النقطة إلى إحدى الحادثتين دون الأخرى أو أن تنتمي إليهما معا، ولا تكون النقطة خارج الاتحاد إلا إذا كانت لا تنتمي إلى A و لا تنتمي إلى B. وهكذا تكون كل نقطة من الاتحاد منتمية إلى واحدة من الحادثتين **على الأقل**، مما يقترح التعريف التالي للاتحاد وهو أيسر وأكثر كفاءة.

(٢ ـ ٥ ـ ٢) اتحاد حادثتين (تعريف آخر)

اتحاد حادثتين هو **حادثة** تتضمن جميع نقاط العينة التي تنتمي إلى واحدة منها على الأقل.

وتتضح كفاءة هذه الصياغة لتعريف الاتحاد من صلاحيته للتعبير عن اتحاد ثلاث حوادث أو أكثر، وفي الحقيقة للتعبير عن اتحاد أي عدد من الحوادث حتى ولو كان لانهائيا فنقول:

(٢ ـ ٥ ـ ٣) اتحاد عدة حوادث

اتحاد n من الحوادث $A_1, A_2, \ldots, A_n$ هو **حادثة** تتضمن كافة نقاط العينة التي تنتمي إلى واحدة منها على الأقل: ونرمز له بـ

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots, \cup A_n$$

(٢ ـ ٥ ـ ٤) تقاطع حادثتين

تقاطع حادثتين A و B هو **حادثة** تتضمن كافة نقاط العينة التي تنتمي إليهما معا. ونرمز له بـ $A \cap B$.

(٢ ـ ٥ ـ ٥) تقاطع عدة حوادث

تقاطع n من الحوادث $A_1, A_2, \ldots, A_n$ هو حادثة تتضمن كافة نقاط العينة التي تنتمي إليها جميعا. ونرمز له بـ

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots, \cap A_n$$

(٢ ـ ٥ ـ ٦) الفرق بين حادثتين

الفرق بين حادثتين A و B هو حادثة تتضمن كل نقاط العينة التي تنتمي إلى A ولا تنتمي إلى B. ونرمز له بـ $A - B$.

(٢ ـ ٥ ـ ٧) تتمـة حادثة

تتمية حادثة A هي حادثة تتضمن كل نقـاط فضاء العينة التي لا تنتمي إلى A. ونرمز لها بـ $\bar{A}$ (أو A^c).

ونـلاحظ أن $\bar{A}$ هـي نفي A، ونعبر عنهـا أحيـانـا بقول «ليس A». كما نلاحظ أن $\bar{A} = S - A$ ، أي الفرق بين فضاء العينة S و A. ومن الواضـح أن الفـرق بين حادثتين A و B هو $A \cap \bar{B}$ ، أي A و ليس B. وذلك من عبارة تعريف الفرق.

(٢ ـ ٥ ـ ٨) الحادثتان المنفصلتان

نقـول إن الحـادثتين منفصلتان إذا كـان تقـاطعهما خاليا، أي $A \cap B = \varnothing$ وتسمى الحادثتان عندئذ متنافيتين.

وهكذا يعنـي تنافي حادثتين أنـه لا يمكن وقوعهما معـا. وهذا واضح مـن عدم وجود أية نقطة عينة مشتركة بينهما. أي أنه لا تـوجد أي نتيجة للتجربة يمكن أن تؤدي إلى تحقق (وقوع) A و B معا. وبعبـارة أخرى، ينفي وقـوع واحدة منهما إمكانيـة وقوع الأخرى في الوقت نفسه.

(٢ ـ ٥ ـ ٩) تجزئة فضاء عينة

نقول إن الحوادث غير المستحيلة (غير الخالية) B_1، B_2، . . . ، B_k تشكل تجزئة لفضاء عينة S إذا حققت الشرطين التاليين:

١) $B_i \cap B_j = \varnothing$ ، $i \neq j$

أي أن الحوادث B_1، B_2، . . . ، B_k متنافيـة مثنى مثنى. (لا يمكن وقوع أي اثنتين منها في وقت واحد).

٢) $\bigcup_{i=1}^{k} B_i = B_1 \cup B_2 \cup \ldots \cup B_k = S$

أي أن اتحاد الحوادث B_1، B_2، . . . ، B_k هو فضاء العينة S. (لابد أن تقع واحدة منها) ونعبر أحيانا عن مثل هذه الحوادث بقولنا إنها متنافية فيما بينها ومُستنفِذة. وبعبارة أخرى، تقع واحدة منها فقط ولابد أن تقع واحدة.

تمارين (٢ ـ ١)

١) نقذف حجر نرد وقطعة نقود، اكتب فضاء العينة S وحدد نقاط العينة في كل من الحوادث التالية:

A : ظهور عدد زوجي على حجر النرد،

B : ظهور وجه الـ H على قطعة النقود،

C : ظهور وجه الـ H على قطعة النقود وعدد أقل من 3 على حجر النرد،

D : ظهور وجه الـ T على قطعة النقود وعدد لا يقل عن 3 على حجر النرد،

E : الحصول على A و B،

F : الحصول على B أو D،

G : الحصول على واحدة على الأقل من الحوادث A، C، D.

من بين الحوادث $\bar{A}$ ، B، C أي الأزواج متنافية؟

٢) قذفنا قطعة نقود ثلاث مرات. اكتب فضاء العينة S ، وعبر عن الحوادث التالية بدلالة نقاط العينة:

A: ظهور وجه الـ H في القذفة الثانية،

B: ظهور وجه الـ H مرتان على الأقل،

C: عدد مرات ظهور وجه الـ H أكبر من عدد مرات ظهور وجه الـ T.

D: وقوع A و $\bar{B}$ ،

E: وقوع A أو C .

٣) اخترنا بذرتين من علبة تتضمن خمس بذور. اثنتان منها تنتجان زهورا بيضاء واثنتان تنتجان زهورا حمراء وواحدة تنتج زهورا زرقاء. اكتب فضاء العينة S .

٤) في الصندوق I كرتان بيضاوان وكرة سوداء، وفي الصندوق II كرة بيضاء وكرة سوداء اخترنا عشوائيا كرة من الصندوق I وخلطناها مع كرات الصندوق II ثم سحبنا منه كرة اكتب فضاء العينة.

٥) نسجل عدد مرات طي سلك نحاسي قبل أن ينقطع. ما هو فضاء العينة.

٦) في خط إنتاج صناعي نسجل عدد القطع التي فحصناها قبل العثور على أول قطعة غير صالحة. ما هو فضاء العينة؟

٧) تقدم شركة خدمات نقل بين مطارين متجاورين، ولديها لهذا الغرض طائرتان مروحيتان تقومان برحلاتهما كل ساعة وعلى مدى الساعات الأربع وعشرين من كل يوم. تحمل الكبرى منهما أربعة ركاب بينما تتسع الصغرى لثلاثة فقط.

ا ـ باستخدام محور إحداثيات بحيث تمثل (x, y) حادثة أنه عند إقلاع الطائرتين في تمام ساعة معينة كانت الكبرى تقل x راكبا بينما يوجد y راكبا على متن الصغرى. ارسم جميع نقاط العينة.

ب ـ صف بكلمات كلا من الحوادث التالية:

$$A = \{(2, 3), (3, 2), (3, 3) (4, 1), (4, 2), (4, 3)\}$$

$$T = \{(0, 0), (1,1), (2, 2),(3, 3)\},$$

$$R = \{(0, 1), (0, 2), (0, 3), (1, 2), (1, 3), (2, 3)\}$$

$$V = \{(0, 3), (1, 2), (2,1), (3, 0)\}$$

جـ ـ اكتب نقاط العينة التي تنتمي إلى كل من المجموعات الجزئية التالية من فضاء العينة وصف بكلمات الحوادث التي تمثلها:

$A \cap T$ ،$T \cup R$ ،$A \cap R$ ،$A \cup V$

د ـ أي الأزواج التالية من المجموعات الجزئية يمثل حادثتين متنافيتين؟

V و A ، R وA ، VوT ، TوR

٨) اخترنا عشوائيا أسرة من مدينة كبيرة ولتكن R حادثة أن الأسرة تمتلك الشقة التي تسكنها، T حادثة أن الأسرة لديها أطفال، و V حادثة أن الأسرة تمتلك سيارة.

بالإشـارة إلى مخطط فن المقابل أذكر (مستخـدما رقم المنطقـة) المنطقة أو المركب من المناطق التي تمثل الحوادث التالية :

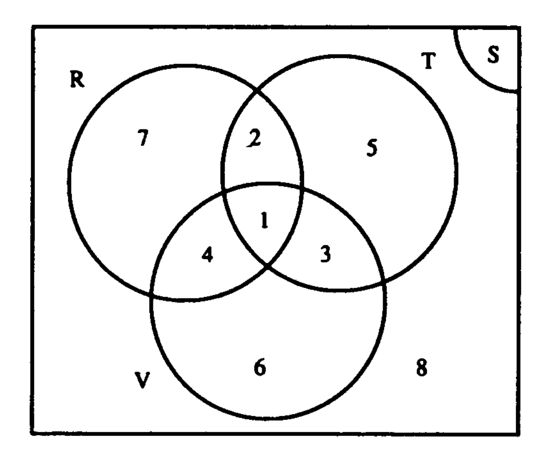
A : الأسرة تمتلك الشقة ولديها أطفال ولا تمتلك سيارة .

B : الأسرة تمتلك الشقة وليس لديها أطفال ولا تمتلك سيارة .

C : الأسرة لا تمتلك الشقة وتملك سيارة .

D : الأسرة لديها أطفال .

E : الأسرة لا تمتلك الشقة وليس لديها أطفال ولا تمتلك سيارة .



٩) بالإشارة إلى التمرين السابق صف بكلمات الحوادث الممثلة بالمناطق التالية :

ا ـ كل منطقـة من المناطق الثماني على حـده . (هل تشكل الحوادث الثماني تجزئة لـ S؟)

ب ـ المنطقة 1 والمنطقة 2 ،

جـ ـ المنطقة 3 والمنطقة 5 ،

د ـ المناطق 3 و 5 و 6 ،

هـ ـ المناطق 1 و 2 و 4 و 7 ،

و ـ المناطق 4 و 6 و 7 و 8 .

١٠) بالاشارة إلى التمرين ٩ عبر عن كل من الحوادث المطلوبة رمزيا بدلالة R ، T ، V .

١١) في المثال (٢ ـ ٢) اكتب الحوادث التالية :

$\bar{A} \cup \bar{B}$ ، $\overline{A \cap B}$ ، $A - B$ ، $A \cup D$ ، $C \cap D$ ، $A \cap B$ ، $A \cup B$.

١٢) في المثال (٢ ـ ٣) اكتب الحوادث التالية :

$\overline{B \cup D}, \bar{B} \cup \bar{D}, \bar{A}, A \cap D, B - D, B \dot{\cup} D, A \cap B$

$A \cap C \cap D, A \cup B \cup D,$

١٣) في المثال (٢ ـ ٤) اكتب الحوادث التالية :

$\bar{T} \cup \bar{B}, T \cap B, T \cup A, \bar{T} \cap \bar{A}, U \cap V, \bar{T}$

هل B و V متنافيتان؟

ملاحظة

من الأمثلة المختلفة التي استعرضناها عن فضاءات العينة نلاحظ أنها إما أن تحوي عددا محدودا (منتهيا) من نقاط العينة، مثل الفضاءات المذكورة في الأمثلة (٢ ـ ١)، (٢ ـ ٢) و (٢ ـ ٣). أو فضاءات تتضمن ما لا نهاية له من نقاط العينة، إلا أنها لا نهاية قابلة للعد، ونقصد بقابلية العد أنه يمكن إقامة تقابل بين نقاط العينة وبين مجموعة الأعداد الطبيعية $\{1, 2, 3, \ldots\}$، ومن الواضح أن وجود هذا التقابل يعني أننا نستطيع عدّ عناصر الفضاء S، فنقول هذا عنصر أول يليه عنصر ثان ثم ثالث ثم رابع وهكذا. . . وهو ما نشاهده في المثال (٢ ـ ٤). ولكن في المثال (٢ ـ ٥) نجد فضاء يتضمن ما لا نهاية له من النقاط، إلا أنها لا نهاية غير قابلة للعد. فالعمر هو قياس زمني يمثل نقطة على محور إحداثي اتخذناه محورا للزمن. ونقاط محور مرصوفة إلى جانب بعضها بصورة متصلة لا انقطاع فيها ولا فجوات. وسواء على كامل المحور أو على أي فترة منه $[a, b]$ ، لا يمكن الإجابة على السؤال التالي: ما العدد أو القياس الذي يلي العدد a مباشرة؟ ومهما حاولنا أخذ عدد قريب من a فسيبقى بينه وبين a ما لا يحصى ولا يعد من القياسات. أي لو أخذنا a عددا أول في محاولة للعد فإنه يستحيل علينا تحديد العدد الثاني. وهكذا نضع اليد على خاصة مميزة لهذا النوع من اللانهايات فنقول إنها لانهاية غير قابلة للعدّ. ويسمى فضاء العينة فضاء منفصلا إذا كانت مجموعة نقاطه منتهية أو لانهائية قابلة للعد. ويسمى فضاء متصلا إذا كانت مجموعة نقاطه لانهائية غير

قابلة للعد. وسنحصل على فضاء متصل من كل تجربة نستخدم فيها، للحصول على النتيجة، جهازا للقياس. وسنحصل على فضاء منفصل في كل تجربة نلجأ فيها، للحصول على النتيجة، إلى عملية تعداد. وستقتصر دراسة الاحتمال في هذا الفصل على فضاءات منتهية أي فضاءات منفصلة تتضمن عددا محدودا من النقاط وسنسميه فضاء منتهيا.

(٢-٦)* أسرة الحوادث - الحقل

تسمى المجموعة التي تكون عناصرها مجموعات صفا أو أسرة. وبدلا من أن نقول مجموعة من المجموعات نقول صفا من المجموعات. إذا عناصر صف أو أسرة هي دائما مجموعات. ولو كتبنا الصف أو الأسرة $\mathscr{A}$ على الشكل:

$$\mathscr{A} = \{A_1, A_2, A_3, \ldots, A_n\}$$

فيجب أن نفهم من هذا أن A_1، A_2، الخ. هي مجموعات من العناصر. وبما أن كل حادثة عبارة عن مجموعة نقاط عينة فسنتحدث عن صف من الحوادث أو أسرة من الحوادث

(٢-٦-١) الحقل

نقول إن أسرة من الحوادث $\mathscr{A}$ تشكل حقلا إذا تحقق الشرطان التاليان:

(١) الأسرة $\mathscr{A}$ مغلقة تحت عملية الاتحاد. (أي أن اتحاد أي حادثتين تنتميان إلى $\mathscr{A}$ هو حادثة تنتمي إلى $\mathscr{A}$ أيضا). وتكتب رمزيا

$$B \in \mathscr{A}\ ,\ A \in \mathscr{A} \Rightarrow A \cup B \in \mathscr{A}$$

مهما تكن A وB من $\mathscr{A}$.

(٢) الأسرة $\mathscr{A}$ مغلقة تحت عملية التتام (أخذ التتمة)..(أي أنه إذا كانت A تنتمي إلى $\mathscr{A}$ فإن $\bar{A}$ تنتمي بدورها إلى $\mathscr{A}$). ونكتب رمزيا

$$A \in \mathscr{A} \Rightarrow \bar{A} \in \mathscr{A}$$

مهما تكن A من $\mathscr{A}$.

---

* للقراءة فقط.

ويمكن البرهان، بسهولة، أن أي حقل من الحوادث يكون مغلقاً تحت عملية التقاطع أي أنه إذا كان $A \in \mathscr{A}$ و $B \in \mathscr{A}$ فإن $A \cap B \in \mathscr{A}$. مهما تكن A و B من $\mathscr{A}$.
ذلك لأن الاستخدام المتتالي لشرطي تعريف الحقل يسمح لنا بالقول:

لتكن A و B أي حادثتين من $\mathscr{A}$ فعندئذ،

$$A \in \mathscr{A} \Rightarrow \bar{A} \in \mathscr{A}$$

$$B \in \mathscr{A} \Rightarrow \bar{B} \in \mathscr{A}$$

ولكن،

$$\bar{A} \in \mathscr{A} \text{ و } \bar{B} \in \mathscr{A} \Rightarrow (\bar{A} \cup \bar{B}) \in \mathscr{A} \Rightarrow \overline{(\bar{A} \cup \bar{B})} \in \mathscr{A}$$

وبما أن

$$\overline{(\bar{A} \cup \bar{B})} = (A \cap B)$$

حسب قانون دي مورغان، فنجد المطلوب.

مثال (٢ ـ ٧)

بالعودة إلى المثال (٢ ـ ٢).

ا ـ اكتب أسرة كافة المجموعات الجزئية من S وتحقّق أنها تشكل حقلا من الحوادث.

ب ـ اكتب أسرة جزئية أو أكثر من أسرة الحوادث المذكورة في ا تُحقِّق شروط الحقل، أي تشكل بدورها حقولا من الحوادث.

الحـــل

ا ـ لنرمز بـ $\mathscr{A}$ لأسرة كل المجموعات الجزية في S فنجد:

$\mathscr{A} = \{\phi, \{(H, H)\}, \{(T, H)\}, \{(H, T)\}, \{(T, T)\}, \{(H, H), (T, H)\}, \{(H, H), (H, T)\}, \{(H, H), (T, T)\}, \{(T, H), (H, T)\}, \{(T, H), (T, T)\}, \{(H, T), (T, T)\}, \{(H, H), (T, H), (H, T)\}, \{(H, H), (T, H), (T, T)\} \{(H, H), (H, T), (T, T)\}, \{(T, H), (H, T), (T, T)\}, S\}$

ومن السهل التحقق من أن اتحاد أي حادثتين من الحوادث الست عشرة التي تتضمنها الأسرة $\mathscr{A}$ ينتمي بدوره إلى $\mathscr{A}$.

ب ـ لنأخذ الأسرة الجزئية $\{\phi, S\}$ فهي تشكل حقلا لأن $\bar{S} = \phi, \ \bar{\phi} = S,$ $\phi \cup S = S$ وشرطا الحقل متحققان.

لنأخذ الآن الأسرة الجزئية التالية ولنرمز لها بـ $\mathscr{F}$:

$$\mathscr{F} = \{\Phi, \{(H, H)\}, \{(H, T)\}, \{(H, H), (H, T)\}, \{(H, T), (T, H), (T, T)\} \{(H, H), (T,H), (T,T)\}, \{(T,H), (T,T)\}, S\}$$

وهي تتضمن ثماني حوادث فقط من $\mathscr{A}$. ومن السهل التحقق من أن اتحاد أي حادثتين من $\mathscr{F}$ ينتمي إلى $\mathscr{F}$. وأن تتمة أي حادثة في $\mathscr{F}$ تنتمي إلى $\mathscr{F}$. فالأسرة $\mathscr{F}$ تشكل حقلا من الحوادث. ويمكن كتابة أسر جزئية أخرى تشكل حقولا. (حاول أن تكتب واحدة).

ملاحظات

١ ـ أسرة كل المجموعات الجزئية من فضاء عينة S. وهي أوسع أسرة حوادث يمكن تشكيلها من S، هي دائما حقل.

٢ ـ كل حقل لابد أن يتضمن فضاء العينة S كأحد عناصره، فهو عندما يتضمن أي حادثة A غير S لابد أن يتضمن تتمة A ، ويتضمن بالتالي $A \cup \bar{A} = S$.

٣ ـ كل حقل لابد أن يتضمن Φ فهو إذ يتضمن S بالضرورة، كما وجدنا في ٢، لابد أن يتضمن تتمة S أي ϕ.

٤ ـ بصورة عامة، يتضمن كل حقل من الحوادث الحادثة المستحيلة Φ والحادثة الأكيدة S . ولو اقتصر الأمر عليهما معا فإنهما يشكلان دائما حقلا. أي أنه من أجل أي فضاء عينة S فإن الأسرة $\{\phi, S\}$ تشكل حقلا.

٥ ـ من أجل أي فضاء عينة S يمكن أن نكتب حقلا أو أكثر من الحوادث في S .

(٢ ـ ٦ ـ ٢) الفضاء الاحتمالي

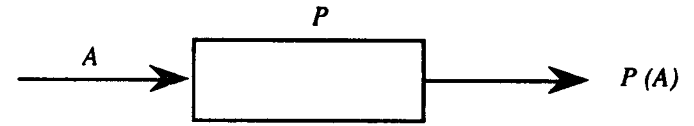
الفضاء الاحتمالي هو ثلاثية $(S, \mathscr{F}, P)$ حيث S فضاء عينة أو الحادثة الأكيدة، $\mathscr{F}$ أسرة من الحوادث في S، P دالة عددية معرفة على الأسرة $\mathscr{F}$ وتحدد لكل حادثة A من الأسرة $\mathscr{F}$ عددا حقيقيا يسمى احتمالها، ونرمز له بـ $P(A)$.

ملاحظـات

١ ـ مسلمات الاحتمال هي حقائق أو أحكام نسلم بصحتها أو بمشروعيتها دون الحاجة إلى برهان. وتشكل الأساس الذي يقوم عليه بناء النظرية الاحتمالية كنظرية رياضية. وتتناول هذه المسلمات الأسرة $\mathcal{F}$ والدالة P. ومعظم الكتاب يقتصرون عند عرض المسلمات على الخواص التي يجب أن تتمتع بها الدالة P، وهو ما سنقوم به في الفقرة القادمة. وتبقى المسلمة المتعلقة بـ $\mathcal{F}$ وكأنها أمر متعارف عليه ضمنا، وسنعلق عليها ونشرح مضمونها هنا في سياق هذه الملاحظات.

هذه المسلمة تقول ببساطة إن الأسرة $\mathcal{F}$ في أي فضاء احتمالي هي حقل. وفي إطار هذه المسلمة فقط يجوز لنا القول إن اتحاد حادثتين هو بدوره حادثة، وأن تتمة حادثة هي الأخرى حادثة، وأن تقاطع حادثتين هو حادثة وهو بالضبط ما تضمنته صياغة التعاريف الواردة في الفقرة (٣ ـ ٥).

٢ ـ يمكن النظر إلى الدالة P وكأنها آلة مصممة من أجل عناصر $\mathcal{F}$ على وجه التحديد. وعندما نُدخل في هذه الآلة عنصرا من $\mathcal{F}$ (أي حادثة) فإنها تخرج لنا عددا هو الاحتمال الموافق.



٣ ـ لدراسة نوع من الظواهر العشوائية احتماليا يكفي تحديد الفضاء الاحتمالي $(S, \mathcal{F}, P)$ الموافق لهذا النوع من الظواهر. وهدف النظرية الاحتمالية هو إقامة مثل هذا الفضاء. ومع تحديد هذا الفضاء يصبح كل ما يهمنا أو يجوز لنا التحدث عن احتماله هو عناصر $\mathcal{F}$. والآلة P مصممة خصيصا لعناصر $\mathcal{F}$ هذه، ولها جميعا دون استثناء وهي تستكمل المهمة المطلوبة فتقدم لنا من أجل كل عنصر من $\mathcal{F}$ (أي من أجل كل حادثة) الاحتمال المقابل.

٤ ـ المسلمة المتعلقة بـ $\mathcal{F}$ والقائلة إن $\mathcal{F}$ حقل تقضي ضمنا ما يلي:
إذا علمنا احتمال وقوع حادثة A فيجب أن نكون قادرين على تحديد احتمال عدم

وقوعها. أليس $\bar{A}$ عنصرا من $\mathscr{F}$؟ إذا P ستقوم بمهمتها في حالة $\bar{A}$) وإذا علمنا احتمال وقوع حادثة A واحتمال وقوع حادثة أخرى B فيجب أن نكون قادرين على تحديد احتمال وقوع A أو B أي احتمال اتحادهما. (أليس $A \cup B$ منتميا إلى $\mathscr{F}$؟ إذا ستقوم الآلة P بمهمتها في حالة $A \cup B$). وكذلك الأمر بالنسبة إلى $A \cap B$.

٥ ـ من الواضح أنه مع الانتهاء من إقامة الفضاء الاحتمالي $(S, \mathscr{F}, P)$ تبقى علينا مهمة لها طابع المهارة التقنية وهي كيفية تشغيل الآلة P لحساب احتمال أي حادثة نريد الحصول على احتمالها. وستكون مهمة القواعد الاحتمالية المختلفة التي نستنبطها هي تصميم آلة P، كفؤة من جهة، وتشغيلها سهل وميسور من جهة أخرى. ويجدر التذكير مجددا أننا نتطرق هنا للفضاءات المنتهية فقط.

(٢ ـ ٧) مسلمات الاحتمال

رأينا أن التكرارات النسبية تسعى إلى الاستقرار بعد سلسلة طويلة من المشاهدات العشوائية التي تجري تحت شروط منتظمة. مما سمح لنا أن ننسب إلى كل حادثة، مرتبطة بتجربة عشوائية، عددا يسمى احتمالها، بحيث أنه عندما نقوم بسلسلة طويلة من التكرارات للتجربة، يصبح التكرار النسبي لوقوع تلك الحادثة مساويا تقريبا لاحتمالها. وقلنا إن هذه هي الصيغة النموذجية للانتظام الاحصائي الذي يشكل الأساس التجريبي لنظرية الاحصاء. كما قلنا إنه عندما نكتشف، عن طريق الملاحظة والتجربة، دلالات كافية على نوع من الانتظام في مجموعة من الظواهر. فإن هذا يدفعنا إلى بلورة نظرية رياضية لمثل هذه الظواهر، تشكل النموذج الرياضي ، أو القالب، الذي يحتوي كافة الحقائق العملية المستوحاة من معطيات الملاحظة والتجربة. وتكون نقطة البداية، عندئذ، هي اختيار أكثر حقائق هذا الانتظام بساطة وجوهرية، وصياغتها في شكل مبسط من جهة، ومجرد ومثالي من جهة أخرى، كموضوعات رياضية نسميها مسلمات، ونعتبرها في مستوى الحقائق المسلم بها سلفا. وسنعرض الآن المسلمات التي تقوم عليها نظرية الاحتمال.

المسلمات

١ ـ $P(A) \geq 0$ ، مهما تكن الحادثة A. (احتمال أي حادثة غير سالب).

٢ ـ $P(S) = 1$ ، حيث S فضاء عينة. (احتمال الحادثة الأكيدة يساوي الواحد).

٣ ـ إذا كانت A_1 ، A_2 حادثتين منفصلتين فإن:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

(احتمال وقوع واحدة منهما على الأقل يساوي مجموع احتماليهما).

تعميم المسلمة الثالثة

ويمكن تعميم المسلمة الثالثة إلى حالة n من الحوادث فنقول:

إذا كانت A_1 ، A_2 ، . . . ، A_n حوادث منفصلة مثنى مثنى فإن:

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n)$$

أو بصورة رمزية مختصرة:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} p(A_i) \quad ; \quad A_i \cap A_j = \phi \quad ; \quad i \neq j$$

ويطلق على المسلمة ٣ اسم الخاصة الجمعية.

والحقيقة أن هذه المسلمات مستوحاة من خواص التكرار النسبي. فإذا كررنا تجربة عشوائية N مرة وراقبنا في كل مرة وقوع أو عدم وقوع حادثة A، مثلا، ورأينا أن A قد وقعت في n من المرات قلنا إن التكرار النسبي لوقوع A كان n/N.

ومن الواضح تماما أن التكرار النسبي لا يمكن أن يكون سالبا. وعندما نقول إن الحادثة A أكيدة فإننا نقصد أن وقوعها محتم في كل مرة نكرر فيها التجربة. أي أن تكرارها النسبي هو الواحد. والمسلمتان الأولى والثانية هما تجريدان لهاتين الحقيقتين التجريبيتين، على الترتيب. وينطبق ذلك أيضا على المسلمة الثالثة. وللايضاح نأخذ المثال التالي:

أحمد طالب في كلية العلوم ـ جامعة الملك سعود يؤدي صلاة الظهر كل يوم في أقرب مسجد لمكان وجوده وقت الظهيرة. لتكن الحادثة A_1 هي أن يصلي أحمد الظهر في

مسجد المبنى ٤ ، A_2 حادثة أن يصلي أحمد الظهر في مسجد المبنى ٥ . سجلنا على مدى ثلاثين يوما تكرار وقوع كل من A_1 و A_2 ووجدنا أن A_1 وقعت عشر مرات A_2 وقعت 8 مرات. فالتكرار النسبي لوقوع A_1 كان $\frac{10}{30}$ ، والتكرار النسبي لوقوع A_2 كان $\frac{8}{30}$. ولو سألنا ما هو التكراو النسبي لحادثة أن يؤدي أحمد صلاة الظهر في المبنى ٤ أو المبنى ٥ لكان الجواب بوضوح $\frac{8}{30} + \frac{10}{30} = \frac{18}{30}$. والتكرار النسبي لوقوع إحدى الحادثتين، على الأقل هو مجموع التكرارين النسبيين لوقوع كل منهما. ونلاحظ أن صحة القاعدة تعود قطعا إلى توفر شرط أساسي هو أنه لا يمكن وقوع A_1 و A_2 في وقت واحد. وفي يوم معين لو رمزنا، مثلا، بـB_1 لحادثة أن أحمد زار المكتبة المركزية، وبـB_2 لحادثة أن أحمد زار مطعم الطلاب. ولاحظنا على مدى ثلاثين يوما أن B_1 وقعت 15 مرة وأن B_2 وقعت عشر مرات، وأنه في خمسة أيام زار كلا من المكتبة والمطعم. فإن التكرار النسبي لوقوع «B_1 أو B_2» ، أي أن يزور أحمد المكتبة أو المطعم، ليس $\frac{25}{30} = \frac{10}{30} + \frac{15}{30}$ لأن الأيام الخمسة التي وقعت فيها كل من B_1 و B_2 حسبناها مرتين، والتكرار النسبي الصحيح هو في الحقيقة $\frac{20}{30} = \frac{15}{30} + \frac{10}{30} - \frac{5}{30}$. ولم نستطع تطبيق القاعدة هنا لأن شرط التطبيق غير متوافر، فالحادثتان B_1 ، B_2 غير منفصلتين، (وقوع إحداهما لا ينفي إمكانية وقوع الأخرى).

(٢ ـ ٨) نتائج

بالاستناد إلى مسلمات الاحتمال يمكننا الآن برهان النتائج التالية

(٢ ـ ٨ ـ ١) إذا كانت A ، B حادثتين بحيث أن $A \subset B$ فإن،

$$P(B-A) = P(B) - P(A)$$

برهان

$$B = BA \cup B\bar{A} = A \cup B\bar{A} \quad ; \quad A \cap B\bar{A} = \phi$$

حسب المسلمة ٣

$$P(B) = P(A \cup B\bar{A}) = P(A) + P(B\bar{A})$$

ومنه :

$$P(B - A) = P(B\bar{A}) = P(B) - P(A)$$

وهو المطلوب .

نتيجة (٢ ـ ٨ ـ ٢)

من أجل أي حادثتين A ، B لدينا :

أ ـ $P(B\bar{A}) = P(B) - P(AB)$

ب ـ $P(A\bar{B}) = P(B) - P(AB)$

برهـــان

من أجل أي حادثتين A ، B لدينا :

$$AB \subset B \quad ; \quad AB \subset A$$

والمطلوب يلي مباشرة من النتيجة السابقة .

نتيجة (٢ ـ ٨ ـ ٣)

إذا كانت A ، B حادثتين وكانت $A \subset B$ فإن $P(A) \le P(B)$.

برهـــان

لدينا من النتيجة ١ ،

$$P(B) = P(A) + P(B - A)$$

ولكن $P(B - A) \ge 0$ حسب المسلمة ١ ، أي أن $P(B)$ لا يمكن أن يكون أقل من $P(A)$ مادام يساوي $P(A)$ مضافا إليه عدد غير سالب .

يمكن التعبير عن النتيجـة (٢ ـ ٨ ـ ٣) بقولنـا إنه كلما اتسعت الحادثـة (أي تضمنت عددا أكبر من نقاط العينة) ازداد احتمال وقوعهـا . أو بعبارة أبسط يزداد احتمال الحادثة كلما اتسعت إمكانيات وقوعها ، أي تعددت الطرق الممكنة التي تؤدي إلى وقوعها . وهو

ما نتوقعه بالفطرة السليمة . وبلغة رياضية تقول النتيجة (٢ ـ ٨ ـ ٣) إن الدالة P ، وتسمى عادة القياس الاحتمالي ، هي دالة غير متناقصة على حقل الحوادث $\mathcal{F}$.

نتيجة (٢ ـ ٨ ـ ٤)

لكل حادثة A لدينا

$$0 \leq P(A) \leq 1$$

برهـــان

لكل حادثة A نعلم أن $A \subseteq S$ وبتطبيق النتيجة (٢ ـ ٨ ـ ٣) والاستفادة من المسلمة ٢ نجد $P(A) \leq P(S) = 1$. أما $P(A) \geq 0$ فيتبع من المسلمة ١ .

نتيجة (٢ ـ ٨ ـ ٥)

$$P(\phi) = 0$$

برهـــان

نعلم أن $\phi \cup S = S$ وأن $\phi \cap S = \phi$

ومنه

$$P(\phi \cup S) = P(S)$$

والطرف الأيسر يساوي $P(\phi) + (P(S)$ حسب المسلمة ٣ ، أي أن

$$P(\phi) + P(S) = P(S)$$

ومن المسلمة ٢ نجد :

$$P(\phi) + 1 = 1$$

ومنه

$$P(\phi) = 0$$

نتيجة (٢ ـ ٨ ـ ٦)

لأي حادثة A لدينا

$$P(A) = 1 - P(\bar{A})$$

برهـــان

لأي حادثة A لدينا $S = A \cup \bar{A}$.

أي أن

$$P(A \cup \bar{A}) = P(S)$$

وبما أن $A \cap \bar{A} = \phi$ نجـد بتطبيق المسلمـة ٣ على الطـرف الأيسر، والاستفـادة من المسلمة ٢ في الطرف الأيمن،

$$P(A) + P(\bar{A}) = 1$$

أو

$$P(A) = 1 - P(\bar{A})$$

نتيجة (٢ ـ ٨ ـ ٧)

لأي حادثة A ، B لدينا :

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

برهـــان

$$A \cup B = A\bar{B} \cup AB \cup \bar{A}B$$

لنفرض أن $P(A\bar{B}) = a$ وأن $P(\bar{A}B) = b$، وأن $P(AB) = c$ ،

فعندئذ

$$P(A \cup B) = P(A\bar{B} \cup AB \cup \bar{A}B) = P(A\bar{B}) + P(AB) + P(\bar{A}B)$$

$$= a + b + c$$

وذلك استنادا إلى المسملة ٣ . ولكن من خواص الأعداد الحقيقية يمكننا كتابة :

$$P(A \cup B) = a + c + b + c - c = (a + c) + (b + c) - c$$

ولكن

$$a + b = P(A\bar{B}) + P(AB) = P(A\bar{B} \cup AB) = P(A)$$

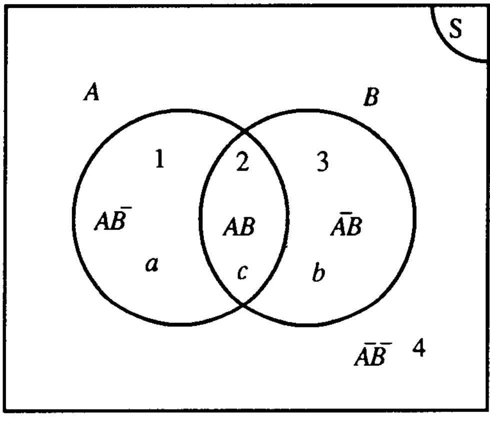
$$b + c = P(\bar{A}B) + P(AB) = P(\bar{A}B \cup AB) = P(B)$$

وذلك بالاستفادة ثانية من المسلمة ٣ . وبالتعويض في العلاقة الأخيرة نجد :

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

ملاحظـة

تتصف قياسات الطول والمساحة والحجم والوزن وسائر القياسات المشابهة بالخاصة الجمعية. إذ لو نظرنا في خريطة مهندس معماري وعليها مخطط لغرفتين متجاورتين، ومساحة أرض الأولى عشرون مترا مربعا ومساحة أرض الثانية ستة عشر مترا مربعا، فاتحاد الغرفتين يعطي غرفة جديدة مساحة أرضها 36 = 20 + 16 مترا مربعا. وكذلك الأمر عند دمج قطعتين منفصلتين من الفضة في قطعة واحدة فوزن القطعة الناتجة هو مجموع وزني القطعتين. والمسلمة الثالثة تقول إن هذه الخاصة الجمعية تبقى صحيحة بالنسبة لاحتمالي حادثتين منفصلتين. وهي تسمح لنا بالنظر إلى احتمالات الحوادث في مخطط ڤن وكأنها مساحات. وبالتالي فإن ما يصح على جمع المساحات نجده صحيحا أيضا على الاحتمالات. وعلى الشكل (٢ ـ ٧) نجد أن المناطق



شكل (٢ ـ ٧)

1، 2، 3، تمثل الحوادث $A\bar{B}$, AB, $\bar{A}B$، على الترتيب. وكما أن مساحة الدائرة A تساوي مساحة المنطقة 1 مضافا إليها مساحة المنطقة 2، فكذلك احتمال الحادثة A يساوي احتمال الحادثة $A\bar{B}$، ممثلة بالمنطقة 1، مضافا إليه احتمال الحادثة AB، ممثلة بالمنطقة 2. ومخطط ڤن في الشكل (٢ ــ ٧) يلعب دور وسيلة الايضاح التي تيسر متابعة وفهم خطوات برهان النتيجة (٢ ـ ٨ ـ ٧) إلا أنه لا يشكل جزءا من البرهان، ولا يجوز أن يكون كذلك.

مثال (٢ ـ ٨)

من أجل أي حادثتين A ، B بين أن :

$$P(A) \leq P(A \cup B),$$
$$P(A) \geq P(A \cap B).$$

نعلم أن

$$A \subseteq A \cup B, \ A \supseteq A \cap B$$

واستنادا إلى النتيجة (٢ ـ ٨ ـ ٣) نجد المطلوب .

مثال (٢ ـ ٩)

بين وجه الخطأ في كل من العبارات التالية :

ا ـ احتمال أن ينجح خالد في امتحان الفيزياء هو 0.95-

ب ـ احتمال أن ينجح خالد في مقرر الاحصاء هو 0.9 واحتمال ألا ينجح هو 0.15 .

جـ ـ احتمال أن يفوز الفريق الوطني لكرة القدم في مباراته القادمة هو 0.75 واحتمال أن يتعادل 0.09 واحتمال أن يفوز أو يتعادل هو 0.95 .

د ـ احتمال أن ينجح خالد في مقرر الاحصاء هو 0.9 واحتمال أن ينجح في مقرري الاحصاء والرياضيات هو 0.95 .

الحـــل

ا ـ يتناقض الاحتمال المعطى مع المسلمة الأولى التي تقول إن احتمال أي حادثة لا يجوز أن يكون سالبا .

ب ـ تناقض الاحتمالات المعطاة المسلمة الثانية . إذ لو رمزنا لحادثة نجاح خالد في مقرر الاحصاء بـA فإن عدم نجاحه يمثل الحادثة المتممة $\bar{A}$ و

$$P(\bar{A}) + P(\bar{A}) = P(S) = 0.9 + 0.15 > 1$$

(احتمال حادثة + احتمال متممتها يجب أن يساوي الواحد بالضبط دون زيادة أو نقصان .)

جـ ـ لنرمز بـ A لحادثة فوز الفريق الوطني لكرة القدم في مباراته القادمة ، ولنرمز بـB لحادثة تعادل الفريق الوطني لكرة القدم في مباراته القادمة .

فيمكن تلخيص المعلومات المعطاة كالتالي :

$$P(A) = 0.75 \,;\, P(B) = 0.09 \,;\, P(A \cup B) = 0.95$$

والحادثتان A ، B متنافيتان وحسب المسلمة الثالثة يجب أن يكون $P(A \cup B)$ مساويا لمجموع $P(A)$ و $P(B)$ وهو غير متحقق لأن $0.95 \neq 0.75 + .09$

د ـ لنرمز بـ A لحادثة أن ينجح خالد في مقرر الاحصاء ،

ولنرمز بـ B لحادثة أن ينجح خالد في مقرر الرياضيات .

لدينا

$$P(AB) = 0.95 \text{ ، } P(A) = 0.9$$

وبما أن

$$AB \subseteq A$$

فلابد أن يكون

$$P(AB) \leq P(A)$$

وفق النتيجة (٢ ـ ٨ ـ ٣) . وهذا غير متوفر، $(0.95 \nleq 0.9)$

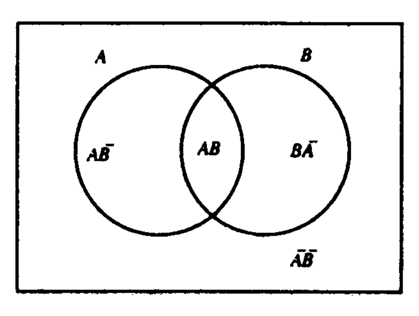
مثال (٢ ـ ١٠)

إذا علمت أن $P(A) = 0.55$ ، $P(A \cup B) = 0.7$ ، $P(B) = 0.35$ فاحسب

$$P(\bar{A}B) \;,\; P(A\bar{B}) \;,\; P(AB)$$

الحـــل

رسم مخطط ڤن مفيد دائما في مثل هذه التمارين . إذ يساعدنا على كتابة العلاقات التي نحتاجها لحل التمرين .



شكل (٣ ـ ٨)

نعلم أن

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

وهي علاقة تربط بين أربعة مقادير. وإذا علمنا أي ثلاثة منها فيمكن استخدامها لحساب المقدار الرابع. لدينا هنا $P(A \cup B)$ و $P(A)$ و $P(B)$ والمطلوب حساب $P(AB)$. بالتعويض في العلاقة نجد

$$0.7 = 0.55 + 0.35 - P(AB)$$

ومنه:

$$P(AB) = 0.55 + 0.35 - 0.7 = 0.2$$

ولدينا أيضا

$$P(A\bar{B}) = P(A) - P(AB)$$
$$= 0.55 - 0.2 = 0.35$$
$$P(\bar{A}B) = P(B) - P(AB)$$
$$= 0.35 - 0.20 = 0.15$$

انظر النتيجة (٢ ـ ٨ ـ ٢).

تمارين (٢ ـ ٢)

١ ـ ما هو وجه الخطأ في كل من العبارات التالية:

ا ـ احتمال هطول مطر في يوم معين هو 0.6 واحتمال وجود رياح نشطة هو 0.8 واحتمال هطول المطر ووجود رياح نشطة هو 0.85.

ب ـ احتمال أن ينجح سالم في مقرر الرياضيات 0.8 واحتمال أن ينجح في مقرر الرياضيات ويرسب في مقرر الفيزياء هو 0.9.

جـ ـ احتمال أن تستقبل عيادة طبيب أقل من 5 مراجعين في فترة ما قبل الظهر هو 0.62 واحتمال أن تستقبل 5 مراجعين أو أكثر هو 0.25.

٢ ـ في دراسة للأحداث الجانحين في مدينة معينة، ترمز R لحادثة أن الجانح ترك المدرسة، وترمز Q لحادثة أن أسرة الجانح ميسورة الحال. أعرض بكلمات الاحتمالات التي تعبر عنها الرموز التالية:

$P(Q' \cup R)$ ، $P(Q'R)$ ، $P(Q'R')$ ، $P(Q \cup R')$ ، $P(QR)$ ، $P(Q')$ ، $P(R')$

٣) إذا كانت D حادثة أن كتابا جديدا في الإحصاء سيُطبع طباعة ممتازة؛ و E حادثة أنه سيلقى رواجا في السوق، و F حادثة أنه سيجري تبنيه لمقرر جامعي. اكتب كلا من الاحتمالات التالية بصورة رمزية:

أ ـ احتمال أن الكتاب سيلقى رواجا ويجري تبنيه لمقرر جامعي.

ب ـ احتمال أن الكتاب سوف لا يُطبع طباعة ممتازة ولا يجري تبنيه لمقرر جامعي.

جـ ـ احتمال أن الكتاب سوف لا يطبع طباعة ممتازة ويجري تبنيه لمقرر جامعي.

د ـ احتمال أن الكتاب سيطبع طباعة ممتازة ويجري تبنيه لمقرر جامعي.

هـ ـ احتمال أن الكتاب سيلقى رواجا ولكنه سوف لا يطبع طباعة ممتازة ولا يجري تبنيه لمقرر جامعي.

٤) بعد تحليل دراسة تمت ضمن كل من ثلا ث شركات يصرح مديروها بما يلي:
يصرح المدير الأول أن احتمالات زيادة في ميزانية الشركة أو انخفاض في ميزانية الشركة، أو بقاء الميزانيةعلى حالتها هي على الترتيب: 0.65 ، 0.07 ، 0.25 .
ويصرح المدير الثاني بأن هذه الاحتمالات بالنسبة إلى شركته هي 0.48، 0.14 ، 0.38،
ويصرح المدير الثالث بأن هذه الاحتمالات بالنسبة إلى شركته هي 0.56 ، 0.08 ، 0.38.
علق على هذه التصريحات من وجهة النظر الاحتمالية.

٥) الحادثتان A و B متنافيتان و $P(A) = 0.12$ ، $P(B) = 0.60$. أوجد:
$P(A')$ ، $P(B')$ ، $P(A \cup B)$ ، $P(AB)$ ، $P(A' \cup B')$ ، $P(A'B')$.

٦) الحادثتان C و D متنافيتان $P(C) = 0.27$ ، $P(D) = 0.33$ ، أوجد:
$P(C')$ ، $P(D')$ ، $P(C \cup D)$ ، $P(C'D)$ ، $P(CD')$ ، $P(C' \cup D')$.

٧) إذا كان 0.2 احتمال أن يبيع معرض سيارات في شهرين ثلاث سيارات على الأقل، احسب احتمال أن يبيع في ذلك الشهر سيارتين على الأكثر.

٨) لنفرض أن $P(A) = 0.56$ ، $P(B) = 0.43$ ، $P(AB) = 0.18$ ، احسب :
$P(A')$ ، $P(B')$ ، $P(A \cup B)$ ، $P(A'B)$ ، $P(AB')$

٩) احتمال أن يحصل مشترك في المسابقة الدولية لتجويد وتفسير القرآن الكريم على جائزة التجويد هو 0.16 واحتمال أن يحصل على جائزة التفسير هو 0.30، واحتمال أن يحصل عليهما معا هو 0.09:
أ ـ احسب احتمال حصول المشترك هذا على واحدة منهما على الأقل .
ب ـ احسب احتمال أن يحصل على واحدة منهما فقط .
جـ ـ احسب احتمال ألا يحصل على أي منهما .

١٠) إذا كان احتمال هطول مطر في يوم معين هو 0.1 ، واحتمال وجود رياح نشطة في ذلك اليوم هو 0.05 واحتمال وجود رياح نشطة وهطول مطر هو 0.03، فاحسب احتمال :
ا ـ هطول مطر أو وجود رياح نشطة في ذلك اليوم ،
ب ـ ألا يهطل مطر في ذلك اليوم ولا توجد رياح نشطة ،
جـ ـ أن توجد رياح نشطة ولا يهطل المطر في ذلك اليوم .

١١) إذا كان $P(B) = 2/3$ ، $P(AB) = 1/2$ ، $P(AB') = 0.25$ فاحسب :
$P(A)$ ، $P(A \cup B)$ ، $P(A'B')$

١٢) إذا علمت أن $P(A) = 0.4$ ، $P(A \cup B)$ ، $P(AB') = 0.25$ فاحسب :
$P(AB)$ ، $P(B)$ ، $P(AB')$ ، $P(A'B')$.

١٣) ما هو وجه الخطأ في كل مما يلي :
ا ـ $P(A) = 0.48$ ، $P(A') = 0.42$.
ب ـ $P(B) = 1.02$ ،
جـ ـ $P(C) - 0.03$ ،
د ـ $P(A) = 0.45$ و $P(AB) = 0.53$ ،
هـ ـ $P(A) = 0.87$ و $P(A \cup B) = 0.79$

١٤) احتمال أن يطلب صاحب سيارة واقف في محطة بنزين الكشف على ضغط الهواء في العجلات هو 0.12 واحتمال أن يطلب الكشف على زيت المحرك هو 0.29، واحتمال أن يطلب الأمرين معا هو 0.07.

أ ـ ما احتمال أن يطلب صاحب سيارة يقف في المحطة الكشف على العجلات أو على زيت المحرك؟

ب ـ ما احتمال أن لا يطلب صاحب سيارة يقف في المحطة الكشف على العجلات ولا الكشف على زيت المحرك؟

جـ ـ ما احتمال أن يطلب صاحب سيارة يقف في المحطة الكشف على العجلات ولا يطلب الكشف على زيت المحرك؟

١٥) يمكن تعميم النتيجة (٢ ـ ٨ ـ ٧) إلى حالة حوادث أو أكثر. بين أن:

$$P(A\cup B\cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

(٢ ـ ٩) بناء نموذج احتمالي

نقصد ببناء النموذج الاحتمالي تحديد طريقة عمل تسمح لنا بحساب قيمة الدالة P لكل حادثة من الحقل $\mathscr{F}$ في فضاء احتمالي $(S, \mathscr{F}, P)$، وبما ينسجم تماما مع المسلمات التي وضعناها.

ومن أجل أي فضاء عينة S، رأينا أنه يمكن تحديد أكثر من حقل من الحوادث، وأبسطها هو الحقل $\{\phi, S\}$، وأكثرها اتساعا هو الحقل المؤلف من كافة المجموعات الجزئية من S. وسنأخذ هذا الحقل بالذات، ولنرمز له فيما يلي بـ $\mathscr{F}$، ونبني عليه نموذجا احتماليا. أي نحدد طريقة ميسرة تقودنا إلى معرفة قيمة الدالة P لأي حادثة من هذا الحقل $\mathscr{F}$. وبالطبع يمكن، بطريقة مماثلة، تعريف نماذج أخرى في حقول أخرى أقل اتساعا.

لنفرض الآن أن عدد نقاط العينة في فضاء عينة S هو t حيث t عدد منته، ولنرمز لهذه النقاط، أو الحوادث الابتدائية، بـ E_1، E_2، . . . ، E_t. من الواضح أن هذه الأسرة

من الحوادث الابتدائية تشكل تجزئة لـS، فهي منفصلة بعضها عن بعض لأن التجربة لا يمكن أن تؤدي إلى نتيجتين مختلفتين في آن واحد، ولأن

$$S = \bigcup_{i=1}^{t} E_i = E_1 \cup E_2 \cup \ldots \cup E_t$$

بمقتضى تعريف فضاء العينة S. وأي حادثة من $\mathcal{F}$ هي، بوضوح، اتحاد عدد من هذه الحوادث الابتدائية المنفصلة.

إذا خصصنا لكل حادثة ابتدائية E_i عددا حقيقيا p_i وكانت هذه الأعداد تحقق الشرطين:

أ ـ $p_i \geq 0$ ، $i = 1, 2, \ldots, t$

ب ـ $\sum_{i=1}^{t} p_i = 1$

نكون قد أقمنا نموذجا احتماليا. إذ نستطيع الآن حساب احتمال أي حادثة من $\mathcal{F}$ كما يلي:

(٢ ـ ٩ ـ ١) احتمال حادثة

احتمال حادثة هو مجموع الاحتمالات المخصصة لنقاط العينة التي تنتمي إلى هذه الحادثة.

وبعبارة أخرى، احتمال حادثة هو مجموع احتمالات الحوادث الابتدائية الداخلة في تشكيل هذه الحادثة.

مثال (٢ ـ ١١)

في المثال (٢ ـ ٤)، افترض أن الاحتمالات المخصصة لنقاط العينة كانت كما في الجدول التالي:

نقطة العينة	(-1,-1)	(-1,0)	(-1,1)	(0,-1)	(0,0)	(0,1)	(1,-1)	(1,0)	(1-1)
الاحتمال المخصص	0.16	0.08	0.16	0.08	0.04	0.08	0.16	0.08	0.16

ا) تحقق أن الجدول يمثل نموذجا احتماليا .

ب) احسب احتمالات الحوادث المذكورة في الجزئين ب و جـ من ذلك المثال .

الحـــل

ا) شرطا النموذج الاحتمالي متحققان، إذ لا يوجد احتمال سالب ومجموع الاحتمالات المخصصة لنقاط العينة يساوي الواحد تماما .

ب)

$$P(A) = P(\{(1,-1)\}) + P(\{(1,0)\}) + P(\{(1,-1)\})$$

$$= 0.16 + 0.08 + 0.16 = 0.40$$

$$P(B) = P(\{(-1,-1)\}) + P(\{(0,0)\}) + P(\{(1,1)\})$$

$$= 0.16 + 0.04 + 0.16 = 0.36$$

$$P(C) = P(\{(-1,\ 0\ \}) + P(\{(1,0)\}) + P(\{(0,-1)\}) + P(\{(0,1)\}) + P(\{(0,0)\})$$

$$= 0.08 + 0.08 + 0.08 + 0.08 + 0.04 = 0.36$$

$$P(U) = P(\{(-1,-1)\}) + P(\{(0,-1)\}) + P(\{(1,-1)\})$$

$$= 0.16 + 0.08 + 0.16 = 0.40$$

$$P(T) = P(\{(-1,1)\}) + P(\{(-1,0)\}) + P(\{(\ 6-1)\}) + P(\{(0,-1) + P(\{(1,-1)\})$$

$$= 0.16 + 0.08 + 0.16 + 0.08 + 0.16 = 0.64$$

$$P(V) = P(\{(-1,1)\}) + P(\{(1,-1)\}) = 0.16 + 0.16 = 0.32$$

مثال (٢ ـ ١٢)

في المثال (٢ ـ ٣) افترض أن حجر النرد هو مكعب متناظر تماما مما لا يترك مبررا لاختلاف الاحتمال المخصص لنقطة عينة من نقطة إلى أخرى من النقاط الست

والثلاثين في فضاء العينة S . [الجدول (٢ ـ ١)]. وفي مثل هذه الحالـة يسمى النموذج «نموذج الاحتمالات المتساوية».

احسب احتمالات الحوادث المذكورة في ب و جـ من ذلك المثال.

الحـــــل

وفقا لنموذج الاحتمالات المتساوية نوزع الـواحد هنا بالتساوي على النقاط الست والثلاثين، فتكون حصة كل منها 1/36. ويكون احتمال أي حادثة في S هو عدد النقاط التي تتضمنها الحادثة مضروبا بـ 1/36. وبذلك تكون الاحتمالات المطلوبة:

$$P(A) = 6 \times \frac{1}{36} = \frac{1}{6} \quad , \quad P(B) = 10 \times \frac{1}{36} = \frac{10}{36} = \frac{5}{18} \; ,$$

$$P(C) = 10 \times \frac{1}{36} = \frac{5}{18} \quad , \quad P(D) = 6 \times \frac{1}{36} = \frac{1}{6} \; ,$$

$$P(E) = 10 \times \frac{1}{36} = \frac{5}{18} \quad , \quad P(F) = P(\phi) = 0 \; ,$$

$$P(G) = 6 \times \frac{1}{36} = \frac{1}{6} \quad , \quad P(H) = 6 \times \frac{1}{36} = \frac{1}{6} \; ,$$

$$P(I) = 4 \times \frac{1}{36} = \frac{1}{9} \quad , \quad P(J) = 6 \times \frac{1}{36} = \frac{1}{6} = \; ,$$

$$P(K) = 9 \times \frac{1}{36} = \frac{1}{4}$$

مثال (٢ ـ ١٣)*

لنفرض في المثال (٢ ـ ٣) أن اهتمامنا يقتصر على المجموع الذي نحصل عليه من القذفتين. أقم على فضاء العينة S نموذجا احتماليا يفي بـالغرض، واستخدمه لحساب احتمالات الحوادث التالية:

T : الحصول على مجموع يساوي 7.

U : الحصول على مجموع يساوي 7 على الأكثر.

V : الحصول على مجموع أكبر من 9.

لنأخذ التجزئة التالية لـ S :

المجموع 2 $\qquad B_1 = \{(1, 1)\}$

---

* للقراءة فقط.

المجموع 3 $B_2 = \{(1, 2), (2, 1)\}$

المجموع 4 $B_3 = \{(1,3), (2,2), (3,1)\}$

المجموع 5 $B_4 = \{(1,4), (2,3), (3,2), (4,1)\}$

المجموع 6 $B_5 = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$

المجموع 7 $B_6 = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

المجموع 8 $B_7 = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$

المجموع 9 $B_8 = \{(3,6), (4,5), (5,4), (6,3)\}$

المجموع 10 $B_9 = \{(4,6), (5,5), (6,4)\}$

المجموع 11 $B_{10} = \{((5,6), (6,5)\}$

المجموع 12 $B_{11} = \{(6,6)\}$

ونلاحظ أن نقاط العينة التي تؤدي إلى المجموع نفسه قد صنفت مع بعضها في المجموعة الجزئية ذاتها. أي أن كل حادثة من حوادث التجزئة B_1 ، B_2 ، . . . ، B_{11} تتضمن نقاط عينة تؤدي إلى المجموع نفسه. وسنخصص احتمالات لهذه الحوادث وفقا للقاعدة الموضحة في المثال السابق انسجاما مع الافتراض بأن حجر النرد مكعب تام التناظر. وبذلك يكون،

$$P(B_1) = \frac{1}{36}\ ,\ P(B_2) = \frac{2}{36}\ ,\ P(B_3) = \frac{3}{36}\ ,\ P(B_4) = \frac{4}{36}\ ,\ P(B_5) = \frac{5}{36}\ ,$$

$$P(B_6) = \frac{6}{36}\ ,\ P(B_7) = \frac{5}{36}\ ,\ P(B_8) = \frac{4}{36}\ ,\ P(B_9) = \frac{3}{36}\ ,\ P(B_{10}) = \frac{2}{36}\ ,$$

$$P(B_{11}) = \frac{1}{36}\ .$$

وسنعرف حقل الحوادث A بأنه الأسرة المؤلفة من حوادث التجزئة B_1 ، B_2 ، . . . ، B_{11} والحوادث الناتجة عن اتحاد أي حادثتين أو أكثر منها بالاضافة إلى ϕ ونعرف الدالة P على هذا الحقل A على الشكل التالي:

احتمال حادثة يساوي مجموع الاحتمالات المخصصة لحوادث التجزئة التي تدخل في تشكيل هذه الحادثة. (لاحظ أن أي حادثة ممكنة يجب أن تكون الآن إحدى حوادث

التجزئة أو اتحاد عدد منها). وبذلك عرفنا الفضاء الاحتمالي $(S, \mathscr{A}, P)$ وحددنا طريقة عمل لحساب الدالة P لكل حادثة من حقل الحوادث $\mathscr{A}$. أي أننا أقمنا نموذجا احتماليا. ومن الواضح أنه نموذج قادر على الاجابة على احتمال أي حادثة تتعلق بالمجموع الذي نحصل عليه في القذفتين. وعلى سبيل المثال:

$$P(T) = P(B_6) = \frac{1}{6}$$

$$P(U) = P(B_1 \cup B_2 \cup B_3 \cup B_4 \cup B_5 \cup B_6)$$

$$= P(B_1) + P(B_2) + P(B_3) + P(B_4) + P(B_5) + P(B_6)$$

$$= \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} = \frac{21}{36} = \frac{7}{21}$$

$$P(V) = P(B_{10} \cup B_{11}) = P(B_{10}) + P(B_{11}) = \frac{2}{36} + \frac{1}{36} = \frac{3}{36} = \frac{1}{12}$$

ويمكن بالطبع استخدام النموذج (الأعم) الذي أقمناه في المثال (٢ ــ ١٢) للحصول على احتمالات الحوادث T ، U ، و V وسنجد الأجوبة نفسها.

تعليق*

على فضاء العينة نفسه المبين في الجدول (٢ ــ ١) ، أقمنا نموذجين احتماليين. وتجدر ملاحظة أن الحقل $\mathscr{F}$ في المثال (٢ ــ ١٢) يتضمن كل المجموعات الجزئية الممكنة من S ، أي 2^{36} حادثة، وهو أوسع حقل يمكن تشكيله من S . والدالة P في المثال (٢ ـ ١٢) تقدم احتمالا لكل من هذه الحوادث. إلا أن الحقل $\mathscr{A}$ في المثال (٢ ـ ١٣) لا يتضمن إلا جزءا يسيرا من حوادث $\mathscr{F}$. والدالة P في المثال (٢ ـ ١٣) تقدم احتمالات الحوادث في $\mathscr{A}$. ولو سألنا مثلا: ما احتمال الحصول على النتيجة نفسها في القذفتين؟ لما أمكن للنموذج المقام في المثال (٢ ـ ١٣) الإجابة عنه، لأن عبارة «الحصول على النتيجة نفسها» ليست حادثة في عُرف الفضاء الاحتمالي $(S, \mathscr{A}, P)$ باعتبارها تمثل مجموعة جزئية غير منتمية إلى الحقل $\mathscr{A}$. وبالتالي ليس لها في عُرف

* للقراءة فقط .

هذا الفضاء احتمال. ولكنها في عُرف الفضاء $(S, \mathscr{F}, P)$ تشكل حادثة لأنها تمثل مجموعة جزئية تنتمي إلى حقل الحوادث $\mathscr{F}$. واحتمالها كما حسبناه في المثال (٢ ـ ١٢) هو 1/6. ولو سألنا في المقابل: ما احتمال الحصول على مجموع زوجي؟ لوجدنا جوابا في النموذج المقام في المثال (٢ ـ ١٣) لأن وصف أو عبارة «المجموع زوجي» يتمثل في $B_1 \cup B_3 \cup B_5 \cup B7 \cup B_9 \cup B_{11}$. أي يمكن التعبير عنه كاتحاد عدد من حوادث التجزئة، وبالتالي فهو ينتمي إلى حقل الحوادث $\mathscr{A}$، أي أنه يمثل حادثة لها في عُرف الفضاء الاحتمالي $(S, \mathscr{A}, P)$ احتمال يساوي

$$P(B_1) + P(B_3) + P(B_5) + P(B_7) + P(B_9) + P(B_{11})$$
$$= \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{7}{36} + \frac{9}{36} + \frac{11}{36} = \frac{18}{36} = \frac{1}{2}$$

ولهذا السؤال جوابه أيضا في النموذج المقام في المثال (٢ ـ ١٢). فالحصول على مجموع زوجي يمثل المجموعة الجزئية المؤلفة من جميع نقاط العينة في الجدول (٢ ـ ١) التي مجموعها زوجي، وعدد هذه النقاط 18. وهي مجموعة جزئية تنتمي إلى $\mathscr{F}$، أي أنها حادثة، وبالتالي لها احتمال يساوي، . وفقا لنموذج المثال (٢ ـ ١٢)، $18 \times 1/36 = 1/2$. وهو الجواب السابق نفسه.

وفي الحقيقة، كل ما يمكن للنموذج في المثال (٢ ـ ١٣) أن يجيب عليه، سيجيب عليه أيضا النموذج «الأوسع» في المثال (٢ ـ ١٢)، ولكن العكس غير صحيح. فالنموذج في المثال (٢ ـ ١٣) صالح للاجابة على حوادث معنية بالمجموع المتحصل من القذفتين فقط. وإذا اقتصر اهتمامنا على مثل هذه الحوادث، فمن الواضح أنه يمكن اعتماد الفضاء $(S, \mathscr{A}, P)$، لأنه يفي بالغرض. وسنرى في الفصل القادم تجسيدا لهذه الفكرة فيما سنسميه بالمتغير العشوائي الذي يولد، اعتمادا على الفضاء الأصلي، فضاء جديدا، لا يُعنى إلا بالقياسات التي يمكن أن يفترضها هذا المتغير العشوائي. وسنسمي النموذج المقام في هذا الفضاء الجديد التوزيع الاحتمالي للمتغير العشوائي.

(٢ ـ ١٠) نموذج الاحتمالات المتساوية

كحالة خاصة لنفرض أن عدد النتائج الممكنة لتجربة هو N، وأنه ليس هناك ما يبرر منح أفضلية لنتيجة من النتائج الممكنة على نتيجة أخرى. فهي جميعها متساوية الأفضلية. أو بعبارة أخرى نقول إن فرصة ظهور نتيجة محددة، عند تنفيذ التجربة، هي نفس فرصة ظهور أي من النتائج الممكنة الأخرى. وقد رأينا مثالا على ذلك في تجربة قذف حجر نرد. وافتراض أن حجر النرد هو مكعب متناظر تماما استدعى القول إن لكل من أوجهه الستة الفرصة نفسها في أن يكون الوجه الظاهر عند قذف الحجر. وفي مثال هذه الحالات نخصص لكل نقطة عينة (نتيجة ممكنة) الاحتمال نفسه، أي نوزع الواحد بالتساوي على النقاط الـ N فتكون حصة كل منها $1/N$. لنفرض الآن أن حادثة A تتضمن n نقطة عينة، فيكون احتمال A حسب التعريف:

$$P(A) = \underbrace{\frac{1}{N} + \frac{1}{N} + \ldots + \frac{1}{N}}_{n \text{ مرة}} = n \times \frac{1}{N} = \frac{n}{N}$$

وهكذا نكون قد أقمنا نموذجا احتماليا يسمى، لأسباب واضحة تماما، نموذج الاحتمالات المتساوية. وفي مثل هذا النموذج يكون احتمال حادثة، باختصار، هو حاصل قسمة عدد النتائج (أو الحالات) الملائمة، على عدد جميع النتائج (أو الحالات) الممكنة، ومنه التعريف التقليدي التالي لاحتمال حادثة:

(٢ ـ ١٠ ـ ١) التعريف التقليدي لاحتمال حادثة

إذا أمكن لتجربة أن تظهر في N من الحالات المتنافية مثنى مثنى والمتساوية الأفضلية. وكان n من هذه الحالات يؤدي إلى تحقق حادثة A فإن احتمال A يساوي n/N.

مثال (٢ ـ ١٤)

في المثال (٢ ـ ٢) إذا افترضنا أن قطعة النقود متناظرة تماما فاحسب احتمالات الحوادث A، B، C، D.

الحـــــل

تناظر القطعة يسمح بتطبيق نموذج الاحتمالات المتساوية أي باستخدام التعريف التقليدي للاحتمال فنجد :

$$P(A) = \frac{\text{عدد الحالات الملائمة}}{\text{عدد الحالات الممكنة}} = \frac{2}{4} = \frac{1}{2}$$

وبصورة مماثلة :

$$P(B) = \frac{2}{4} = \frac{1}{2} \; , \; P(C) = \frac{3}{4} \; , \; P(D) = \frac{3}{4}$$

مثال (٢ ـ ١٥)

يتضمن صندوق أول كرتين بيضاويين وكرة سوداء واحدة . ويتضمن صندوق ثان كرة بيضاء وكرة سوداء . سحبنا عشوائيا كرة من الصندوق الأول وخلطناها جيدا مع كرات الصندوق الثاني ، ثم سحبنا عشوائيا كرة من الصندوق الثاني . ما هو احتمال الحصول على كرة بيضاء؟

تمييزا للكرات ذات اللون نفسه بعضها عن بعض نرقمها فنضع الرقم 1 على إحدى الكرتين البيضاوين في الصندوق الأول ، ولنرمز لها بـW_1 ، والرقم 2 على الكرة البيضاء الأخرى في الصندوق الأول ، ولنرمز لها بـW_2 ، والرقم 3 على الكرة البيضاء في الصندوق الثاني والرقم 1 على الكرة السوداء في الصندوق الأول ، ولنرمز لها بـB_1 ، والرقم 2 على الكرة السوداء في الصندوق الثاني ، ولنرمز لها بـB_2. ويمكن الآن كتابة فضاء العينة S كما يلي :

$$S = \{ W_1 W_1 , \; W_1 W_3 , \; W_1 B_2 , \; W_2 W_2 , \; W_2 W_3 , \; W_2 B_2 , \; B_1 B_1 , \; B_1 W_3 , \; B_1 B_2 \}$$

حيث ترمز نقطة العينة $W_2 W_3$ ، مثلا ، إلى النتيجة : «سحبنا الكرة W_2 من الصندوق الأول . ثم سحبنا الكرة W_3 من الصندوق الثاني» . ويتضمن S تسع نقاط . والسحب

العشوائي من كل من الصندوقين يعني أن لكل من النتـائج التسع الفرصة نفسها في أن تكون النتيجة التي نحصل عليها عند تنفيذ التجربة. وحصة كل نقطة عينة هي إذا 1/9.
والحادثة المطلوبة، ولنرمز لها بـA ، تتضمن النقاط التالية:

$$A = \{W_1\ W_1,\ W_1\ W_3,\ W_2\ W_2,\ W_2\ W_3,\ B_1\ W_3\}$$

ويكون:

$$P(A) = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{5}{9}$$

أو نطبق التعريف التقليدي لـلاحتمال فنقوم بتعـداد النتائج الملائمة وهـي النقاط التي يكون حرفها الثاني W فنجدها 5 ويكون:

$$P(A) = \frac{\text{عدد الحالات الملائمة}}{\text{عدد الحالات الممكنة}} = \frac{5}{9}$$

مثال (٢ ـ ١٦)

لدينـا خمس بذور، اثنتان منها تنتجـان زهورا حمراء ولنرمـز لها بـR_1 وR_2. واثنتان تنتجان زهورا بيضاء، ولنرمز لهما بـ W_1، W_2 ، وواحدة تنتج زهورا صفراء، ولنرمز لها بـy. خلطنا هذه البذور جيدا ثم اخترنا منهـا عشوائيا بذرتين. فما هو احتمال أن تنتجا زهورا من اللون نفسه؟

فضاء العينة هو:

		الاختيار الثاني				
		R_1	R_2	W_1	W_2	Y
الاختيار الأول	R_1	–	$R_1 R_2$	$R_1 W_1$	$R_1 W_2$	$R_1 Y$
	R_2	$R_2 R_1$	–	$R_2 W_1$	$R_2 W_2$	$R_2 Y$
	W_1	$W_1 R_1$	$W_2 R_2$	–	$W_2 W_1$	$W_1 Y$
	W_2	$W_2 R_1$	$W_2 R_2$	$W_1 W_2$	–	$W_2 Y$
	Y	YR_1	YR_2	YW_1	YW_2	–

وهو يتضمن عشرين نقطة عينة، لكل منها فرصة 1/20 في أن تكون هي النتيجة التي يتمخض عنها الاختيار. وعدد النتائج التـي تحقق المطلوب، هو عدد النقاط التي تتضمن الحرف نفسه، وهو 4 . فالاحتمال المطلوب يساوي 1/5=4/20.

أو كان يمكن الاكتفاء بعشر «حالات» تتضمن كل منها نقطتين لهما، إذا أغفلنا ترتيب الحرفين فيهما، المدلـول نفسه. فمثلا، W_1 R_1 وR_1 W_1 تعنيان في النـاتج النهائي الحصول على بذرتين هما W_1 و R_1 دون أخذ الترتيـب الذي حصلنـا فيه على W_1 أو R_1 في الاعتبار وطالما أن كلا من الحالات العشر تتضمن نقطتي عينـة فلهما أفضليات متساوية وفرصة كل منهما هي 1/10. ومن بين هذه الحالات العشر الممكنة نجـد حالتين ملائمتين فقط والاحتمال المطلوب يساوي 1/5=2/10، وهو الجواب نفسه الذي حصلنـا عليه منذ قليل.

وقد يقـول قائل لماذا لا نختصر إلى حـالتين فقط فإما أن نحصل على زهـرتين من اللـون نفسـه أو تكـون الـزهرتـان من لـونين مختلفين؟ فهناك حـالتان ممكنتـان أحدهما ملائمة والأخرى غير ملائمة والجواب حسب التعريف هو 1/2. وهو جواب مختلف عن الجوابين السابقين المتساويين، وخاطىء طبعا لأن أحـد شرطي التعريف (٢ ـ ٢٠ ـ ١) غير متــوفر. فالحالتـان هنـا متنافيتـان فعلا ولكن فـرصة إحـداهما 4/20، بينما فرصـة الأخرى 16/20. (تتضمن ست عشرة نقطة عينة) أي أن شرط الأفضليـات المتساوية غير متوفر.

سـؤال

بالعودة إلى المثال (٥ ـ ١٣) لنعتبر كل حادثـة من حوادث التجزئة حالة، فيكون لدينا إحدى عشرة حالة ممكنة. ولحساب احتمال الحادثة V نجد ثلاث حالات ملائمة ويكون الجواب 3/11 وهو جواب خاطىء. لماذا؟

(٢ ـ ١١) الاحتمال الاحصائي

قُذفت قطعة نقود، تبدو متوازنة ومتناظرة، مئة مرة، وسُجلت النتائج في الجدول (٢ ـ ٢) ، حيث سجلنـا التكرار النسبي لظهـور كل من وجهي الـ H والـT. وكما رأينا

في الفقرة (٢ ـ ٢) فإن التكرار النسبي سيميل إلى الاستقرار حول قيمة محددة عندما نستمر في تكرار التجربة عددا كبيرا من المرات. وفي الجدول (٢ ـ ٢) نجد أن التكرار النسبي قريب من 1/2، وهذا ليس مفاجئا، فتناظر قطعة النقود سيجعلنا نتوقع ظهور وجه الـH حوالي نفس عدد مرات ظهور وجه الـT.

جدول (٢ ـ ٢)

النتيجة	التكرار	التكرار النسبي الملحوظ	التكرار النسبي المتوقع على المدى المطويل من قطعة متزنة
H	56	0.56	0.50
T	44	0.44	0.50
المجموع	100	1.00	1.00

وفي تجربة أخرى، قُذف حجر نرد، يبدو متناظرا، 300 مرة، وسجل تكرارظهور كل من الأوجه الستة. فكانت النتائج كما في الجدول (٢ ـ ٣). ونلاحظ اقتراب التكرار النسبي لظهور كل من الأوجه الستة من القيمة 1/6. وهذه النتائج غير مفاجئة بدورها، طالما أن حجر النرد يبدو متناظرا ومتزنا. مما يقترح علينا اعتبار التكرار النسبي في الجدول (٢ ـ ٢)، تقديرا أوليا لاحتمال ظهور وجه معين من وجهي قطعة النقود عند قذفها، واعتبار التكرار النسبي، في الجدول (٢ ـ ٣)، لظهور وجه معين من أوجه حجر النرد تقريبا لاحتمال ظهور ذلك الوجه عند قذف حجر النرد.

وفي الحقيقة، يمكننا، كما رأينا في الفقرة (٢ ـ ٢)، أن نفترض وجود عدد p هو احتمال وجه الـH. وإذا بدت لنا القطعة متوازنة ومتناظرة تماما فيمكن بطريقة استنتاجية القول إن احتمال كل وجه هو 1/2. أما إذا لم تكن القطعة تامة التناظر، كما هو الحال في الواقع العملي، فيمكن اللجوء إلى التجربة، فنقذف القطعة عددا كافيا من المرات، ونسجل النتائج كما في الجدول (٢ ـ ٢)، ثم نعتبر التكرار النسبي لوجه الـH كتقريب لقيمة p.

جدول (٢ ـ ٣)

النتيجة	التكرار	التكرار النسبي	التكرار النسبي المتوقع على المدى الطويل من حجر نرد متزن
1	51	0.170	0.1667
2	54	0.180	0.1667
3	48	0.160	0.1667
4	51	0.170	0.1667
5	49	0.163	0.1667
6	47	0.157	0.1667
المجموع	300	1.000	1.0000

وفي حالة حجر النرد يمكننا، عند افتراض تناظر الحجر تماما، القول بعدم وجود أفضلية لوجه على الآخر، وإن المنطق يدعو إلى الاستنتاج بأن احتمالات ظهور كل من الأوجه الستة ولنرمز لها بـ p_1، p_2، . . . ، p_6 متساوية وكل منها يساوي 1/6. ولما كان الحجر المتناظر تماما غير موجود إلا في مخيلتنا، ولا يمكن الوصول إلى صناعة حجر نرد متناظر تماما. إلا أنه يمكن أن تكون صناعة الحجر متقنة فيبدو لنا وكأنه متناظر تماما. وعندئذ سنستمر في اعتبار 1/6 قيمة تقريبية جيدة لكل من p_1، p_2، . . . ، p_6. ولو فرضنا الآن أن حجر النرد غير متوازن، وأنه من المؤكد أن أوجهه الستة لا تتمتع بفرص الظهور نفسها عند قذف الحجر، ففي هذه الحالة لا يزال ممكنا بالطبع افتراض وجود الأعداد p_1، p_2، . . . ، p_6، إلا أنه لا يمكن تقدير أي منها بطريقة استنتاجية. ولا بد من اللجوء إلى التجربة فنقذف الحجر عددا كبيرا من المرات ثم نعتبر التكرار النسبي لظهور كل وجه تقديرا لاحتمال ظهور ذلك الوجه.

وكما رأينا في مطلع هذا الفصل، فإن معظم الظواهر التي نواجهها في حياتنا العملية، هي من النوع الذي لا يمكن التنبؤ بنتائجه سلفا. فلنفرض، مثلا، أننا نريد تقدير احتمال أن يكون أول طفل سيولد في مدينة الرياض ذكرا. مثل هذه الحادثة

تصادفية، ويمكننا، استنادا إلى ظاهرة الانتظام الاحصائي، أن نفترض وجود عدد p يسمى احتمالها. ولا يمكن، في الواقع العملي، معرفة p تماما، إلا أنه يمكن تقديرها بصورة جيدة. ولو عدنا، مثلا، إلى سجلات الولادات في مدينة الرياض لفترة سنوات خلت، فوجدنا أن 51% من الولادات كانت ذكورا، فيكون معقولا أن نعتبر 0.51 قيمة تقريبية لـ p. والاحتمال الذي نحصل عليه بهذه الطريقة يسمى أحيانا الاحتمال الاحصائي.

تمارين (٢ ـ ٣)

١) يتضمن صندوق ست قطع حمراء من الورق مرقمة من 1 إلى 6، وكذلك ست قطع بيضاء من الورق مرقمة من 1 إلى 6. وجميع القطع من الحجم نفسه. سحبنا قطعة بصورة عشوائية. ما احتمال أن تكون:
ا ـ حمراء؟ ب ـ عليها رقم زوجي؟ جـ ـ حمراء وعليها رقم زوجي؟
د ـ حمراء أو عليها رقم زوجي؟ هـ ـ ليست حمراء وليس عليها رقم زوجي؟

٢) قذفنا حجري نرد. لتكن A حادثة الحصول على عدد فردي من القطعة الأولى، وB حادثة الحصول على عدد أكبر من 2 من القطعة الثانية.
احسب $P(A)$، $P(B)$، $P(AB)$، $P(A \cup B)$.

٣) إذا كانت احتمالات أن تتلقى عيادة طبيب 0، 1، 2، 3، 4، 5، 6، 7 أو أكثر من المكالمات الهاتفية خلال ساعة الظهر هي، على الترتيب، 0.001، 0.006، 022.، 052.، 0.091، 0.128، 0.149، 0.551. فما هو احتمال أنها ستتلقى:
ا ـ أقل من 5 مكالمات؟ ب ـ ثلاث مكالمات على الأقل؟
جـ ـ من 2 إلى 4 مكالمات؟

٤) احتمالات تقويم هيئة المواصفات والمقاييس لأداة مبتكرة للوقاية من التلوث بغاز السيارات، بأنها رديئة، مقبولة، جيدة، ممتازة، هي على الترتيب: 0.12، 0.23،

0.45، 0.20. احسب احتمال أن يكون تقويمها للأداة:
ا ـ رديئة أو مقبولة، ب ـ على الأقل مقبولة،
جـ ـ جيدة أو ممتازة، د ـ مقبولة أو جيدة.

٥) يعلم صاحب مطعم من خبرته السابقة أن احتمالات أن يطلب زبون بعد تناول الغداء، بوظة، معمولا، فطاير بالجوز، فطاير بالقشطة، عصير برتقال، بطيخا، هي، على الترتيب، 0.13، 0.24، 0.09، 0.11، 0.05، 0.07. ما هو احتمال أن يطلب زبون ما يلي:
أ ـ بوظة أو عصير برتقال؟
ب ـ فطاير بالجوز أو فطاير بالقشطة أو معمول؟
جـ ـ بوظة أو معمولا أو عصير برتقال أو بطيخا؟
د ـ لا شيء مما ذكر؟
علما أن الزبون يقدم رغبة واحدة فقط.

٦) في التمرين ٧ من مجموعة التمارين (٢ ــ ١) لنفرض أن لكل نقطة من فضاء العينة الاحتمال نفسه (نموذج الاحتمالات المتساوية) احسب احتمالات الحوادث التالية:
أ ـ A، T، R، V.
ب ـ احسب احتمالات الحوادث الواردة في الجزء جـ من ذلك التمرين.

٧) بالاشارة إلى التمرين ٢ أعلاه، لتكن C حادثة الحصول على مجموع زوجي. احسب $P(C)$، $P(AC)$، $P(BC)$، $P(A \cup C)$، $P(ABC)$، $P(A \cup B \cup C)$

٨) في التمرين ١ من مجموعة التمارين (٢ ـ ١). إذا فرضنا أن حجر النرد وقطعة النقود يتصفان بالتناظر التام.
احسب احتمالات الحوادث A، B، C، D، E، F، G.

٩) في التمرين ٢ من مجموعة التمارين (٢ ــ ١). احسب احتمالات الحوادث A، B، C، D، E، بفرض أن قطعة النقود متناظرة.

١٠) في السوق عـرض مخفض لبيع مجموعة من المعلبات التي لا عنـوان عليها . ويحوي هذا البيع 200 علبة طماطم، 300 علبة سبانخ، 100 علبة مشمش، و400 علبة كمثرى، فما احتمال أن أول مبتـاع سيحصل على علبـة خضروات؟ علبة فـواكه؟ علبة كمثرى؟

١١) في مسح صحي تناول عينة ضخمة من السكان في بلد معين تم تشخيص الاصابة أو عدم الاصابة بالديدان . وكانت النتائج كما هو مبين في الجدول التالي :

نسبة المصابين بالديدان في هذه الشريحة من العمر	النسبة من العينة في هذه الشريحة من العمر	شريحة العمر (بالسنوات)
0.09	0.20	0 - 4
0.25	0.18	5 - 9
0.31	0.14	10 - 14
0.62	0.09	15 - 19
0.49	0.13	20 - 25
0.41	0.10	30 - 39
0.41	0.07	40 - 49
0.40	0.04	50 - 59
0.28	0.05	60 +

إذا اخترنـا عشوائيـا شخصا من هـذه العينة السكـانية فما احتمال أن يكـون (أو أن تكون) :

ا ـ من 15 سنة إلى 19 سنة؟

ب ـ أقل من 15 سنة؟

جـ ـ من 15 سنة إلى 29 سنة؟

د ـ من 15 إلى 19 ومصاب بالديدان؟

هـ ـ من 15 إلى 29 ومصاب بالديدان

و ـ من 15 إلى 29 وغير مصاب بالديدان؟
ز ـ مصاب بالديدان؟
ح ـ ما هو احتمال أن شخصا من 15 إلى 29 مصاب بالديدان؟

١٢) لأغراض عدة يُقال إن الطفل خديج إذا كان وزنه عند الولادة 5.5 باوند أو أقل مستخدما البيان الاحصائي المعطى في التمرين ١٥ من مجموعة التمارين (١ ـ ١)، احسب:
أ ـ احتمال أن طفلا مولودا عام 1965 في جنوب غرب انكلترا مسجل خديجا.
ب ـ الوزن عند الولادة الذي يجري تخطيه باحتمال 0.025
جـ ـ الوزن عند الولادة الذي يجري تخطيه باحتمال 0.975.

(٢ ـ ١٢) طرق العد

في المثالين (٢ ـ ١٥) و (٢ ـ ١٦) استطعنا بجهد مقبول وضع قائمة تتضمن كافة نقاط العينة. ولكن ماذا لو أن عدد النتائج الممكنة كان كبيرا جدا؟ لا شك أن اللجوء إلى حصر النتائج واحدة فأخرى سيكون شاقا، وغالبا ما يكون من الناحية العملية مستحيلا. وسنستعرض الآن عددا من القواعد المفيدة والسهلة التي يمكن استخدامها للوصول بسرعة ويسر إلى عدد الحالات الممكنة وعدد الحالات الملائمة التي وردت في التعريف التقليدي لاحتمال حادثة.

(٢ ـ ١٢ ـ ١) قاعدة الـ $m \times n$

إذا أمكن استكمال مرحلة أولى من عمل معين بـ m طريقة، ومن أجل كل من هذه الطرق أمكن لمرحلة ثانية أن تتم بـ n طريقة، فالعدد الكلي للأشكال المختلفة لاستكمال العمل بمرحلتيه هو $m \times n$ طريقة.

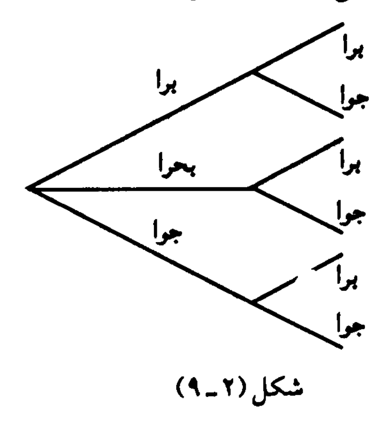
مثال (٢ ـ ١٧)

يمكن لحاج أن يصل جدة برا أو جوا أو بحرا وبعد إتمام مناسك الحج يمكنه الوصول إلى المدينة المنورة برا أو جوا. فبكم طريقة مختلفة يمكن لحاج إتمام مناسك الحج وزيارة المسجد النبوي الشريف؟

المرحلـة الأولى يمكن أن تتم بثـلاث طـرق ، ومـن أجل كل منهـا يمكن أن تتم المرحلة ا لثانية بطريقتين ، فعدد الطرق المختلفة الممكنة ،

$$3 \times 2 = 6$$

ويوضح المخطط في الشكل (٢ ـ ٩) الطرق الست الممكنة .



شكل (٢ ـ ٩)

مثال (٢ ـ ١٨)

بكم طريقة يمكن كتابة زوج مرتب عنصره الأول أحد الأعداد 1,2,3,4,5,6 وعنصره الثاني أحد الحروف a,b,c,d,e؟

الجواب $6 \times 5 = 30$

ويبين الجدول (٢ ـ ٤) الأزواج المرتبة الثلاثين بالتفصيل .

جدول (٢ ـ ٤)

	1	2	3	4	5	6
a	(a,1)	(a,2)	(a,3)	(a,4)	(a,5)	(a,b)
b	(b,1)	(b,2)	(b,3)	(b,4)	(b,5)	(b,6)
c	(c,1)	(c,2)	(c,3)	(c,4)	(c,5)	(c,6)
d	(d,1)	(d,2)	(d,3)	(d,4)	(d,5)	(d,6)
e	(e,1)	(e,2)	(e,3)	(e,4)	(e,5)	(e,6)

ويمكن تعميم قاعدة الـ $m \times n$ إلى عمل يتضمن k من المراحل المتتالية. ولو فرضنا أنه يمكن إتمام الرحلة الأولى بـ n_1 طريقة والمرحلة الثانية بـ n_2 طريقة، . . . ، والمرحلة الـ k بـ n_k طريقة فيكون عدد الطرق المختلفة لاتمام العمل بجميع مراحله هو

$$n_1 \times n_2 \times \ldots \times n_k$$

(٢ ـ ١٢ ـ ٢) المتبادلات

يسمى ترتيب r من الأشياء المتميزة «متبادلة». لنفرض أن لدينا n شيئا متميزا ونريد اختيار r شيئا منها ثم ترتيبها في متبادلة، فبكم طريقة مختلفة يمكن القيام بذلك؟

ونرمز عادة لعدد الطرق هذا بـ P_r^n ويقرأ «عدد متبادلات n شيئا مأخوذ r منها في وقت واحد».

نظرية المتبادلات

$$P_r^n = n(n-1)\ldots(n-r+1)$$

برهان

المسألة المطروحة مكافئة لمسألة شغل r من المواقع المحددة المتتالية وذلك بأن نضع في كل موقع شيئا نختاره من بين الأشياء الـ n المتوفرة. ومثل هذا العمل يتضمن بوضوح r مرحلة. فالمرحلة الأولى شغل الموقع الأول، والمرحلة الثانية شغل الموقع الثاني وهكذا. . . والمرحلة الأخيرة شغل الموقع الـ r. ويمكن، بوضوح، شغل الموقع الأول بأي شيء نختاره من بين الأشياء الـ n المتوفرة، أي بـ n طريقة مختلفة. ويمكن شغل الموقع الثاني باختيار أي شيء من الأشياء الـ $n-1$ المتبقية، أي بـ $n-1$ طريقة مختلفة وهكذا، . . . ، والموقع الأخير يمكن شغله بـ $n-(r-1) = n-r+1$ طريقة مختلفة. ووفقا لقاعدة الـ $m \times n$ المعممة نجد المطلوب.

مثال (٢ ـ ١٩)

اشتريت مرجعا من خمسة أجزاء. وعلى رف من رفوف مكتبتك في المنزل لا يتوفر إلا ثلاثة أمكنة. بكم طريقة مختلفة يمكنك شغل هذه الأماكن الثلاثة المتوفرة بثلاثة أجزاء تختارها من الأجزاء الخمسة؟

الحـــل

عدد الطرق المختلفة لشغل الأماكن الثلاثة هو عدد متبادلات خمسة أشياء مأخوذ ثلاث منهـا في وقت واحد أي P_3^5 . ولحساب P_3^5 نطبق نظريـة المتبادلات ، فنحسب القوس الأخيرة $(n - r + 1)$حيث $n = 5$ ، $r = 3$ لنجد

$$n - r + 1 = 5 - 3 + 1 = 3$$

ويكون P_3^5 مساويا لجداء الأعداد الصحيحة المتناقصة بدءا من 5 وانتهاء بـ 3 ، أي

$$P_3^5 = 5 \times 4 \times 3 = 60$$

وفي الحالة الخاصة التي يكون فيها $r = n$ ، نطبق نظرية المتبادلات بوضع $r = n$ ، نجد أن عدد متبادلات n شيئا مأخوذة جميعها في وقت واحد هو

$$P_n^n = n(n-1)(n-2) \times \ldots \times 2 \times 1$$

ونرمز لمثل هذا الجداء بـ $n!$ ويُقرأ «مضروب n ».

وهذا يعني أن عدد الطـرق المختلفة لترتيب n شيئا متميزًا هو $n!$. وباستخدام رمز المضروب يمكن التعبير عن P_r^n كما يلي :

$$P_r^n = n(n-1) \ldots (n-r+1)$$

$$= \frac{[n(n-1) \ldots (n-r+1)] \, [(n-r)\,(n-r-1) \ldots \times 2 \times 1]}{[(n-r)(n-r-1) \ldots \times 2 \times 1]}$$

$$= \frac{n!}{(n-r)!}$$

ولو عوضنا r بـ n لوجدنا

$$P_n^n = \frac{n!}{0!}$$

ويفتـقر مضروب الصفـر، $(0!)$، إلى أي مغـزى عملي ، ولكن رأينـا قبـل قليـل أن $P_n^n = n!$ وهذا يؤدي إلى أنه لابد أن يكون $n! = \frac{n!}{0!}$ ، مما يقترح علينا أن نصطلح على اعتبار $0!$ مساويا للواحد $(0! = 1)$.

(٢ ـ ١٢ ـ ٣) المتوافقات

إذا كان لدينا مجموعة تتضمن n عنصرا فاختيار مجموعة جزئية من r عنصرا ($r \leq n$) يسمى متوافقة. وعدد المجموعات الجزئية المختلفة التي يمكن اختيارها يسمى «عدد متوافقات n شيئا مأخوذ r منها في وقت واحد». ونرمز له عادة بـ C_r^n أو $\binom{n}{r}$، ونقرؤها «n متوافقات r» أو «n اختيار r». وتجدر هنا ملاحظة أن لا أهمية لترتيب اختيار العناصر. فالمجموعة الجزئية من r عنصرا نختارها من بين n عنصرا ستبقى بدون تغيير طالما تضمنت العناصر نفسها، وذلك بصرف النظر عن الترتيب الذي تم فيه اختيار هذه العناصر.

ومن الواضح أنه يجب أن تكون هناك علاقة بين P_r^n، حيث نختار «اختيارا مرتباً، وC_r^n إذ لا أهمية لترتيب الاختيار. وللوصول إلى هذه العلاقة نحاول حساب P_r^n بالطريقة التالية، فنقول إنه يمكن الوصول إلى P_r^n على مرحلتين، حيث نختار في المرحلة الأولى جميع متوافقات n شيئا مأخوذ r منها في وقت واحد، ولنرمز لعدد هذه المتوافقات بـ C_r^n كما أسلفنا، ثم نرتب عناصر كل متوافقة فور اختيارها بجميع الأشكال الممكنة، ونعلم أن عدد مثل هذه الترتيبات المختلفة أو المتبادلات يساوي $r!$. والعملية هنا تتألف إذا من مرحلتين، أولاهما يمكن أن تتم بـ C_r^n طريقة، ومن أجل كل من هذه الطرق يمكن أن تتم المرحلة الثانية بـ $r!$ طريقة. وحسب قاعدة الـ $m \times n$ يمكن إتمام العملية المطلوبة بـ $C_r^n \times r!$. وهذا يعني أن

$$P_r^n = C_r^n \times r!$$

أو

$$C_r^n = \frac{P_r^n}{r!} = \frac{n!}{r!(n-r)!}$$

وبذلك نكون قد برهنا النظرية التالية:

نظرية المتوافقات

عدد متوافقات n شيئا مأخوذ r منها في وقت واحد هو:

$$C_r^n = \frac{n!}{r!(n-r)!}$$

مثال (٢ ـ ٢٠)

في المثال (٢ ـ ١٩) بكم طريقة يمكنك اختيار ثلاثة منها لوضعها على رف المكتبة؟

الحل

يقتصر المطلوب على اختيار ثلاثة أجزاء من بين خمسة دون أهمية للترتيب الذي حصل فيه الاختيار. فالاختيار سيكون نفسه، مثلا، إذا بدأنا باختيار الجزء الخامس ثم اخترنا بعده الثالث ثم الأول، أو بدأنا باختيار الجزء الأول ثم اخترنا بعده الثالث ثم ختمنا بالخامس، . . . ، وهكذا يمكن أن نمضي فنذكر ستة ترتيبات مختلفة لاختيار هذه الأجزاء بعينها، هي على وجه التحديد:

$$(1,3,5); (1,5,3); (3,1,5); (3,5,1); (5,3,1); (5,1,3)$$

ومن حيث مضمون الاختيار (وهو ما يقتصر عليه اهتمامنا في المتوافقات) فإن الترتيبات الستة تؤدي إلى الاختيار نفسه، أو إلى المتوافقة نفسها. وهذا يوضح أن كل ست متبادلات قد اختزلت إلى متوافقة واحدة. وبذلك يكون العدد المطلوب هو

$$\frac{P_3^5}{6} = \frac{60}{6} = 10$$

وسنجد الجواب نفسه بتطبيق نظرية المتوافقات، السابقة فنكتب:

$$C_3^5 = \frac{5!}{3!\,(5-3)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{(3 \times 2 \times 1) \times 2!}$$
$$= \frac{5 \times 4}{2!} = \frac{20}{2 \times 1} = 10$$

ملاحظة

يمثل C_r^n، كما رأينا، عدد المجموعات الجزئية من r عنصرا التي يمكن اختيارها من مجموعة تتضمن n عنصرا. ولكن اختيار مجموعة جزئية من r عنصرا يعني عمليا تقسيم المجموعة التي نختار منها إلى مجموعتين، إحداهما تتضمن الـ r عنصرا التي اختيرت، والأخرى تتضمن الـ $n - r$ عنصرا المتبقية. وبالتالي فإن $C_r^n = \frac{n!}{r!\,(n-r)!}$ تجيب على سؤال آخر يمكن صياغته على الشكل التالي:

بكم طريقة يمكن تقسيم n شيئا متميزا إلى قسمين أحدهما يتضمن n_1 شيئا والآخر يتضمن n_2 شيئا، حيث $n_1 + n_2 = n$؟

والجواب هو

$$C^{n}_{n_1} = C^{n}_{n_2} = \frac{n!}{n_1!\ n_2!}$$

والغاية من طرح المسألة بهذه الصيغة هي قابليتها للتعميم بسهولة. فبكم طريقة يمكن تقسيم n شيئا متميزا إلى ثلاثة أقسام أولها يتضمن n_1 شيئا والثاني يتضن n_2 شيئا والثالث يتضمن n_3 شيئا حيث $n_1 + n_2 + n_3 = n$؟ والجواب ببساطة هو

$$\frac{n!}{n_1!\ n_2!\ n_3!}$$

ويمكن برهان ذلك بأن نقوم بعملية التقسيم المطلوب على مرحلتين. فنقسم الأشياء المتميزة الـ n إلى قسمين أحدهما يتضمن n_1 شيئا والآخر يتضمن الـ $n_2 + n_3$ شيئا المتبقية. ثم نقوم في المرحلة الثانية بتقسيم الـ $n_2 + n_3$ شيئا إلى قسمين أحدهما يتضمن n_2 شيئا والآخر يتضمن n_3 شيئا. ومما سبق نعلم أن عدد الطرق المختلفة لاتمام المرحلة الأولى هو $\frac{n!}{n_1!(n_2+n_3)!}$. وعدد الطرق المختلفة لاتمام المرحلة الثانية هو $\frac{(n_2+n_3)!}{n_2!\ n_3!}$ وعدد الطرق المختلفة لاتمام عملية التقسيم بمرحلتيها هو حسب قاعدة الـ $m \times n$:

$$\frac{n!}{n_1!(n_2+n_3)!} \times \frac{(n_2+n_3)!}{n_2!\ n_3!} = \frac{n!}{n_1!\ n_2!\ n_3!}$$

ويمكن تعميم هذه النتيجة فنقول أن عدد طرق تقسيم n شيئا متميزا إلى k قسما، يتضمن القسم الأول n_1 شيئا منها، ويتضمن القسم الثاني n_2 شيئا وهكذا، . . .، ويتضمن الجزء الأخير n_k شيئا، حيث $n_1 + n_2 + \cdots + n_k = n$، هو

$$\frac{n!}{n_1!\ n_2!\ \ldots\ n_k!}$$

(٢ ـ ١٢ ـ ٤) متبادلات n من الأشياء غير المتميزة

وللقاعدة التي توصلنا إليها في ختام الملاحظة السابقة تطبيق هام. فلنفرض أن لدينا n من الأشياء غير المتميزة، حيث n_1 منها أشياء متطابقة ومن النوع نفسه، و n_2

منها متطابقة ومن النوع نفسه، وهكذا . . . ، وn_k منها متطابقة ومن النوع نفسه، فبكم طريقة مختلفة يمكن ترتيب هذه الأشياء؟ أي ما هو عدد متبادلات الأشياء الـn مأخوذة جميعها في وقت واحد؟

لو عدنا إلى تصور عملية الترتيب كعملية مكافئة لوضع الأشياء الـn في n من المواقع المتتالية لأمكننا أن نقول ما يلي:

سنحصل على متبادلة لهذه الأشياء الـn عندما نقسم المواقع الـn إلى k قسما، أولها يتضمن n_1 موقعا نضع فيها أشياء النوع الأول، وثانيها يتضمن n_2 موقعا نشغلها بأشياء النوع الثاني، وهكذا . . . ، وآخرها يتضمن الـn_k موقعا الباقية لتأوي إليها أشياء النوع الأخير. وإذ لا تتغير المتبادلة عندما يتبادل شيئان من النوع نفسه موقعيهما، فإنها تتغير في حالة واحدة فقط وهي عندما نقوم بنقل شيء من نوع معين إلى موقع شيء من نوع آخر. وعدد المتبادلات المختلفة هو إذا عدد الطرق المختلفة لعملية التقسيم تلك، أي

$$\frac{n!}{n_1!\ n_2!\ \dots\ n_k!}$$

(٢ ـ ٢١) مثال

ما عدد متبادلات حروف كلمة Statistics؟

تتضمن الكلمة عشرة حروف. ويتكرر الحرف s ثلاث مرات والحرف t ثلاث مرات، والحرف a مرة، والحرف i مرتين، والحرف c مرة واحدة.

الحـــل

$$\text{عدد المتبادلات} = \frac{10!}{3!\ 3!\ 2!\ 1!\ 1!} = 50400$$

مثال (٢ ـ ٢٢)

تريد هيئة للرقابة والتفتيش تشكيل ثلاث لجان لدراسة موضوع الأسعار في صناعة معيّنة. ويتوافر عندها 84 مفتشا. فبكم طريقة يمكن تشكيل اللجان الثلاث إذا كانت ستتضمن 17، 19و 27 مفتشا. وأنه لا يمكن لمفتش أن يشترك في أكثر من لجنة واحدة؟

الحــــل

العدد المطلوب هو عدد إمكانات تقسيم الـ 84 مفتشا إلى أربع مجموعات إحداها تتضمن 17 مفتشا، والثانية تتضمن 19 مفتشا، والثالثة تتضمن 27 مفتشا، والرابعة تتضمن الـ 21 مفتشا الباقين. أي

$$\frac{84!}{17!\ 19!\ 27!\ 21!}$$

مثال (٢ ـ ٢٣)

من حقيبة تحوي 7 كرات سود و 5 كرات بيض، سحبنا عشوائيا خمس كرات فما احتمال أن تتضمن كرتين بيضاوين؟

الحــــل

عدد الحالات الممكنة هو C_5^{12} . وعدد الحالات الملائمة هو عدد طرق اختيار كرتين بيضاوين من الكرات الخمس البيض، مضروبا بعدد طرق اختيار الكرات السود الباقية من بين الكرات السود السبع المتوفرة. أي $C_2^5 \times C_3^7$

ويكون الاحتمال المطلوب:

$$\frac{\left(C_2^5 \times C_3^7\right)}{C_5^{12}} = \frac{5!}{2!\ 3!} \times \frac{7!}{3!\ 4!} + \frac{12!}{5!\ 7!}$$

$$= \frac{4 \times 5}{2} \times \frac{5 \times 6 \times 7}{2 \times 3} + \frac{8 \times 9 \times 10 \times 11 \times 12}{5 \times 4 \times 3 \times 2}$$

$$= (10 \times 35) + 792 = \frac{350}{792} = 0.442$$

مثال (٢ ـ ٢٤)

في المثال (٢ ـ ١٦) احسب الاحتمال المطلوب بتطبيق نموذج الاحتمالات المتساوية.

الحــــل

عدد الحالات الممكنة هو $C_2^5 = 10$. وعدد الحالات الملائمة هو بوضوح اثنان. البذرتان اللتان تنتجان زهورا حمرا أو البذرتان اللتان تنتجان زهورا بيضا) ويكون الاحتمال المطلوب

$$\frac{2}{10}=0.2$$

تمارين (٢ ـ ٤)

١) قاعة للاحتفالات فيها أربعة أبواب. بكم طريقة مختلفة يمكنك الدخول إلى القاعة والخروج منها؟

٢) يذاكر أحمد كل يوم إما 0 أو 1 أو 2 ساعة. بكم طريقة يمكن لأحمد أن يذاكر ما مجموعه أربع ساعات في ثلاثة أيام متتالية؟

٣) توجد أربعة طرق A، B، C، D بين منزلك والجامعة. إذا كان للطريق A اتجاه واحد هو من الجامعة إلى المنزل وللطريق D اتجاه واحد هو من المنزل إلى الجامعة.
ا ـ ارسم رسما توضيحيا يبين عدد الامكانات المختلفة للقيام برحلتك اليومية إلى الجامعة ذهابا وإيابا.
ب ـ شريطة أن يختلف طريقا الذهاب والاياب كم يصبح عدد الامكانات المختلفة للقيام برحلتك؟

٤) بكم طريقة مختلفة يمكنك ترتيب خمسة من كتبك الجامعية على رف مكتبتك؟

٥) بكم طريقة يمكن تشكيل عدد من أربعة أرقام:
ا ـ إذا كان التكرار ممكنا؟
ب ـ إذا لم يكن التكرار ممكنا؟

٦) ما عدد أرقام الهواتف الممكنة المؤلفة من سبعة منازل عشرية إذا كانت المنزلة الأخيرة 3 أو 4؟

٧) إذا توافر عشرة لاعبين لكرة سلة فكم فريقا من خمسة لاعبين يمكن تشكيله إذا أمكن لكل لاعب أن يقوم بأي دور يوكل إليه؟

٨) توجد ستة مواضيع تعبير مختلفة يختار الطالب في ١٠١ نجل واحدا منها للكتابة فيه. فبكم طريقة يمكن لأربعة طلاب في هذا المقرر أن يختاروا مواضيعهم بحيث:
أ ـ لا يختار طالبان الموضوع نفسه.
ب ـ لا توجد أية قيود على اختيار المواضيع.

٩) بكم طريقة يمكن لمدير محطة تليفيزيون أن يوزع ستة إعلانات تجارية على ستة أوقات مخصصة للدعاية أثناء إذاعة مباراة في كرة القدم؟

١٠) فصل يتضمن عشرين طالبا، منهم 15 من المستوى الأول، و5 من المستوى الثاني. بكم طريقة يمكن اختيار لجنة من ثلاثة طلاب بحيث:
ا ـ تتضمن واحدا من المستوى الثاني واثنين من المستوى الأول؟
ب ـ تتضمن واحدا على الأقل من المستوى الثاني؟

١١) مجموعة من خمس عشرة ساعة فيها ساعة واحدة معيبة. بكم طريقة يمكن أن نختار منها ثلاث ساعات بحيث:
أ ـ لا تتضمن الساعة المعيبة؟
ب ـ تتضمن الساعة المعيبة؟

١٢) بالاشارة إلى التمرين السابق لنفرض أن المجموعة تتضمن ساعتين معيبتين، فبكم طريقة يمكن اختيار ثلاث ساعات بحيث تكون:
ا ـ جميعها سليمة؟
ب ـ واحدة منها معيبة؟

١٣) تحقق أن

$$\binom{n}{r}+\binom{n}{r-1}=\binom{n+1}{r} \quad , \quad \binom{n}{r}=\binom{n}{n-r}$$

١٤) بالاشارة إلى التمرين ٩، بكم طريقة يمكن للمدير شغل ستة أوقات للاعلانات التجارية إذا كان لديه أربعة إعلانات مختلفة ويتكرر أحدها ثلاث مرات؟

١٥) ما هو عدد التباديل المختلفة لحروف كلمة INDEPENDENCE؟

١٦) يتضمن اختبار «صح ـ خطأ» ستة عشرة سؤالا. احسب عدد الطرق المختلفة لإعداد ورقة الإجابة. احسب عدد الطرق التي يمكن أن نختار فيها:
ا ـ ثمانية أسئلة للإجابة عليها بـ «صح» وثمانية للإجابة عليها ب «خطأ».
ب ـ عشرة أسئلة للإجابة عليها بـ «صح» وستة للإجابة عليها بـ «خطأ».

١٧) توجد في متاهة أربعة تقاطعات. وعند كل منها يمكن لفأر أن يذهب يمينا أو يسارا أو على خط مستقيم. ما احتمال اجتياز الفأر للمتاهة عند أول محاولة، إذا علمت أنه يوجد طريق واحد ممكن بين طرفي المتاهة؟

١٨) قدمنا لقرد اثنتي عشرة قطعة تتضمن ثلاثة مربعات، وثلاثة مستطيلات، وثلاثة مثلثات، وثلاث دوائر. إذا رتب بنجاح ثلاثا من الشكل نفسه، ثم ثلاثا من شكل ثان، ثم ثلاثا من شكل ثالث، وثلاثا من الشكل الرابع المتبقي. ما احتمال هذه الحادثة تحت الفرض بأن القرد لا يميز بالفعل بين الأشكال الهندسية؟
١٩) بالاشارة إلى التمرين ٦، ما احتمال أن يكون رقم هاتفك 4343434؟

٢٠) بالاشارة إلى التمرين ١٠، إذا اخترنا لجنة بصورة عشوائية فما هو احتمال أن تتضمن واحدا على الأقل من المستوى الثاني؟

٢١) بالاشارة إلى التمرين ١١، لنفرض أننا اخترنا عشوائيا ثلاث ساعات فما احتمال أن تتضمن الساعة المعيبة؟

٢٢) بالاشارة إلى التمرين ١٢، لنفرض أننا اخترنا عشوائيا ثلاث ساعات فما احتمال أن تكون جميعها سليمة؟

(٢ ـ ١٣) الاحتمال الشرطي

عندما تستيقظ صبيحة يوم من أيام فصل الشتاء، وتنظر إلى السماء لتجدها ملبدة بالغيوم، فسيكون احتمال هطول المطر في ذلك اليوم أقوى مما لو وجدت سحابا متفرقا. ولو رمزنا لحادثة «هطول المطر» بـ A، ولحادثة «السماء ملبدة بالغيوم» بـ B. ورمزنا بـ $P(A|B)$ لاحتمال A علما أن B قد وقعت، أي احتمال هطول المطر علما بأن السماء ملبدة بالغيوم، فإن $P(A|B)$ سيكون أكبر من $P(A)$، وهذا بدوره أكبر من $P(A|\bar{B})$. فمعرفتنا المسبقة بأن السماء غائمة، تعني أن الفرصة مهيأة بمشيئة الله لسقوط المطر، مما يزيد من احتمال A. ويخفض هذا الاحتمال معرفتنا المسبقة بأن السماء صافية. ويسمى $P(A|B)$ الاحتمال الشرطي لـ A علما أن B قد وقعت.

ويوضح هذا المثال أن الحوادث قد تكون، بصورة عامة، على صلة ببعضها، بمعنى أن وقوع حادثة قد يؤثر زيادة أو نقصانا في احتمال وقوع حادثة أخرى. ومن هنا تأتي أهمية الاحتمال الشرطي. ولو وجدنا أن وقوع B لم يؤثر لا زيادة ولا نقصانا في احتمال وقوع A، أي أن $P(A|B) = P(A)$، فسنستنتج بلا شك أن لا صلة للحادثتين ببعضهما من الناحية الاحتمالية، أو أنهما مستقلتان احتماليا. وسنتعرض لمفهوم الاستقلال في فقرة قادمة.

مثال (٢ ـ ٢٥)

قذفنا حجر نرد متوازن، ولتكن:

A_1: حادثة الحصول على 2،

A_2: حادثة الحصول على عدد أقل من 4،

A_3: حادثة الحصول على عدد أقل من 5،

B: حادثة الحصول على عدد زوجي.

احسب $P(A_1|B)$، $P(A_2|B)$، $P(A_3|B)$

الحـــل

يتضمن فضاء العينة ست نقاط. وطالما أن الحجر متوازن فلا توجد أفضلية لوجه

على آخر، وحصة كل نقطة عينة هي 1/6، كما هو مـوضح في الشكل (٢ ـ ١٠). ومن السهل رؤية أن:

$$P(A_1) = \frac{1}{6} \quad , \quad P(A_2) = \frac{1}{2}$$
$$P(A_3) = \frac{2}{3} \quad , \quad P(B) = \frac{1}{2}$$

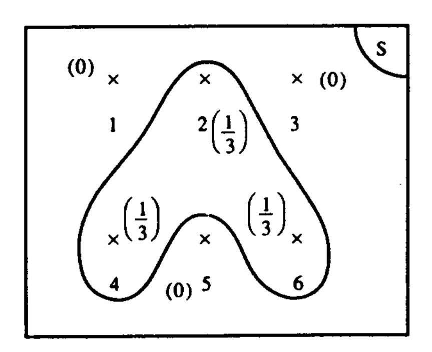
			S
$(\frac{1}{6})$ × 1	$(\frac{1}{6})$ × 2	$(\frac{1}{6})$ × 3	
$(\frac{1}{6})$ × 4	$(\frac{1}{6})$ × 5	$(\frac{1}{6})$ × 6	

شكل (٢ ـ ١٠): النموذج غير الشرطي

لنفرض الآن أن الشخص الـذي قذف حجر النرد أفادنـا أن النتيجة التي حصل عليها كانت زوجية، أي أن الحادثة B قد وقعت.

ولنستعرض آثار هـذه المعلومات، التي توفـرت لنا مسبقا، على الاحتمالات التي حسبناها أعلاه دون أي شروط مسبقة. فنقول أولا إنه لابد من بناء نموذج احتمالي جديد يأخذ في الاعتبـار حقيقة أن B قد وقعت، وأن $P(B)$ الآن هو الواحـد. وفي ظل هذه الحقيقة لم تعد النتائج الممكنة ستة، وإنما أصبحت ثلاثا فقط. فهي إما 2 أو 4 أو 6. أما النتائج 1، 3، 5 فأصبحت مستحيلة. ولا يجوز عنـد بناء النموذج الجديـد أن نمنحها حصة غير الصفر. ونحن هنا أمام فضاء جديد يسمى الفضاء الشرطي، وإذ استخدمنا الحرف P رمزا لـدالة الاحتمال في الفضاء غير الشرطي، فمن المستحسن استخدام الـرمز P_B لدالـة الاحتمال في الفضاء الشرطـي. وهي تذكـرنا أن الاحتمالات محسـوبة الآن على أساس أن الحادثة B قد وقعت. ولكن ما هي الاحتمالات التي تخصصها الدالة P_B لكل من نقاط العينة الستة؟ من الواضح أولا أن

$$P_B(\{1\}) = P_B(\{3\}) = P_B(\{5\}) = 0$$



شكل (٢ ـ ١١): الفضاء الشرطي والنموذج المقام عليه

ومجموع الاحتمالات أو الحصص التي كانت الدالة P تمنحها لهذه النقاط ، ويساوي النصف ، يجب أن توزعه P_B على النقاط 2، 4، 6. فكيف تتم عملية التوزيع هذه؟ من الواضح أن كل نقطة من هذه النقاط ينبغي أن يزداد احتمالها بصورة تتناسب طردا مع الاحتمال الذي خصصته لها الدالة P. ولو أن P خصصت لنقطة ω_i ، مثلا ، ضعف ما خصصته لنقطة أخرى ، ω_j ، فإن حصة ω_i من الزيادة ينبغي لها أن تكون ضعف حصة ω_j منها . وفي مثالنا هنا حيث خصصت P احتمالات متساوية لكل من 2 ، 4 ، 6 ينبغي أن توزع P_B النصف المتوفر بالتساوي على هذه النقاط ليصبح الاحتمال الجديد لكل منها 1/3 .

أي
$$P_B(\{2\}) = P_B(\{4\}) = P_B(\{6\}) = \frac{1}{3}$$

وهكذا تقيم P_B على فضاء العينة S نموذجا جديدا هو النموذج الشرطي ، [انظر الشكل (٢ـ١١)] ويكون :

$$P(A_1|B) = P_B(A_1) = P_B(\{2\}) = \frac{1}{3} > P(A_1) ,$$

$$P(A_2|B) = P_B(A_2) = P_B(\{2\}) = \frac{1}{3} < P(A_1) ,$$

$$P(A_3|B) = P_B(A_3) = P_B(\{2, 4\}) = P_B(\{2\}) + P_B(\{4\}) ,$$

$$= \frac{1}{3} + \frac{1}{3} = \frac{2}{3} = P(A_3)$$

تغيرت النتـائج في ضوء الحقيقة التي عرفـناها (حقيقة وقوع B)، فزاد احتمال A_1 من 1/6 إلى 1/3 ، وانخفض احتمال A_2 من 1/2 إلى 1/3 ، أما احتمال A_3 فلم يتغير.

وقـد لا تكـون إقـامـة النموذج الشرطـي الجديد الـذي نستخـدمـه في حسـاب الاحتمالات الشرطية، عمـلا سهلا. وسنقدم الآن تعـريفا لـلاحتمال الشرطي يسمح لنا بـاستخـدام النمـوذج غير الشرطي لحساب الاحتمالات الشرطيـة بيسر وسهـولـة، دون الحاجة إلى كتابة أو ذكر الفضاء الشرطي والنموذج المقام عليه.

تعريف الاحتمال الشرطي

لتكن A ، B حادثتين في فضاء عينة S فعندئذ:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0 ;$$

أو

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) \neq 0 .$$

وهذا التعريف يقول ببساطة: لحساب الاحتمال الشرطي لحادثة A علما أن حادثة أخرى B قـد وقعت، نقسم احتمال وقـوع A و B معـا على احتمال وقـوع B فنجد المطلوب.

لنعد الآن إلى المثال السابق ولنحسب:

$$P(A_1 B) = P(\{2\}) = \frac{1}{6}$$

$$P(A_1|B) = \frac{P(A_1 B)}{P(B)} = \frac{1/6}{1/2} = \frac{2}{6} = \frac{1}{3}$$

وبصورة مماثلة:

$$P(A_2 B) = P(\{2\}) = \frac{1}{6}$$

$$P(A_2|B) = \frac{P(A_2 B)}{P(B)} = \frac{1/6}{1/2} = \frac{2}{6} = \frac{1}{3}$$

و

$$P(A_3 B) = P(\{2\}) + P(\{4\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$P(A_3 | B) = \frac{P(A_3 B)}{P(B)} = \frac{1/3}{1/2} = \frac{2}{3}$$

وهي النتائج ذاتها التي وصلنا إليها باستخدام الفضاء الشرطي. إلا أننا في جميع الحسابات هنا لم نحتج حتى إلى التفكير بالفضاء الشرطي، ولم نستخدمه.

مثال (٢ ـ ٢٦)

صنّفنا مائة شخص وفقا للجنس (ذكر، أنثى) ووفقا للإصابة بمرض عمى الألوان (مصاب، غير مصاب). فكانت النتيجة كما في الجدول التالي:

	مصاب	غير مصاب	المجموع
ذكر	2	58	60
أنثى	1	39	40
المجموع	3	97	100

اخترنا عشوائيا شخصا واحدا ولتكن:

A حادثة الشخص مصاب بعمى الألوان،

B حادثة الشخص ذكر.

إذا علمنا أن الشخص الذي تم اختياره كان ذكرا فما هو احتمال أن يكون مصابا؟ نعلم الآن أن الاختيار كان من 60 ذكرا بينهم اثنان من المصابين فالاحتمال المطلوب هو

$$P(A | B) = \frac{2}{60}$$

وبصورة مماثلة، إذا علمنا أن الشخص الذي اختير مصاب، فاحتمال كونه ذكرا، هو نسبة الذكور بين المصابين، والاختيار كان من ثلاثة مصابين، بينهم اثنان من الذكور، والاحتمال المطلوب هو:

$$P(B | A) = \frac{2}{3}$$

ولو حسبنا $P(A)$ ، $P(B)$، و$P(AB)$ ، ثم طبقنا التعريف لوجدنا :

$$P(AB) = \frac{2}{100} \quad , \quad P(B) = \frac{60}{100} \quad , \quad P(A) = \frac{3}{100}$$

ومنه :

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{2/100}{60/100} = \frac{2}{60} ,$$

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{2/100}{3/100} = \frac{2}{3} .$$

وهي الأجوبة السابقة نفسها .

مثال (٢ ـ ٢٧)

أظهر تصنيف لطلاب الجامعة أن 10% من الطلاب يدخنون ، وأن 30% من الطلاب يشربون القهوة ، وأن 5% من الطلاب يدخنون ويشربون القهوة .

أ ـ احسب النسبة المئوية للطلاب الذين لا يدخنون ولا يشربون القهوة .
ب ـ من بين الطلاب المدخنين ما هي نسبة الطلاب الذي يشربون القهوة؟
جـ ـ من بين الطلاب الذين لا يشربون القهوة ما هي نسبة المدخنين؟

الحــل

نلاحظ ، بصورة عامة ، أنه إذا كان لدينا مجتمع فيه N عنصرا ، ومن بينهم n عنصرا يتصف بصفة معينة C ، مثلا ، فإن نسبة العناصر في هذا المجتمع التي تتصف بالصفة C هي n/N. أو، كنسبة مئوية ، نقول إن n/N 100 بالمائة من هذا المجتمع يتصفون بالصفة C. ولكن n/N هي بالضبط احتمال أن نختار عشوائيا عنصرا من هذا المجتمع فنجده يتصف بالصفة C. (عدد الحالات الملائمة مقسوما على عدد الحالات الممكنة) . أي أن احتمال أن نختار، بصورة عشوائية ، عنصرا واحدا من هذا المجتمع فنجده متصفا بالصفة C هو ببساطة نسبة الذين يتصفون بالصفة C في المجتمع . وهذا يوضح كيف نترجم النسبة إلى احتمال وكيف نفسر الاحتمال كنسبة . الأمر الذي وجدنا مبرراته في الفقرات (٢ـ٢) ، (٢ـ١٠) و (٢ـ١١) .

لنتصور أن التجربة هي اختيار عشوائي لطالب من طلاب الجامعة ولنرمز بـ
A لحادثة الطالب يدخن .
B لحادثة الطالب يشرب القهوة .
أ ـ لحساب النسبة المطلوبة نحسب احتمال الحادثة AB ثم نفسره كنسبة . ولكن (حسب قانون دي مورغان والنتيجة ٢ ـ ٨ ـ ٦)

$$\begin{aligned} P(\bar{A}\bar{B}) &= 1 - P(A \cup B) \\ &= 1 - [P(A) + P(B) - P(AB)] \\ &= 1 - 0.10 - 0.30 + 0.05 = 0.65 \end{aligned}$$

أي أن 65% من الطلاب لا يدخنون ولا يشربون القهوة .

ب ـ لحساب هذه النسبة التي تشترط أن الطالب مدخن نحسب $P(B|A)$ ثم نفسره كنسبة .

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{0.05}{0.10} = \frac{1}{2}$$

أي أن 50% من الطلاب المدخنين يشربون القهوة .

جـ ـ ولحساب هذه النسبة حيث نشترط أن الطالب لا يشرب القهوة . نحسب $P(A|\bar{B})$ ثم نفسره كنسبة .

$$P(A|\bar{B}) = \frac{P(A\bar{B})}{P(\bar{B})} = \frac{P(A) - P(AB)}{1 - P(B)} = \frac{0.10 - 0.05}{1 - 0.30} = 0.071$$

أي أن 7.1% فقط من الطلاب الذين لا يشربون القهوة مدخنون .

تعليق*

نقدم فيما يلي برهاناً للعلاقة الواردة في تعريف الاحتمال الشرطي ، حيث نرمز لنقطة عينة بـ ω. ولاحتمال حادثة A علماً أن الحادثة B قد وقعت بـ $P_B(A)$. وبـ $P(A)$ لاحتمال A غير الشرطي . ونعلم أولاً أن :

* للقراءة فقط .

$$\sum_{\omega\in B} P_B(\{\omega\}) = 1$$

حيث نقصد بالرمز $\sum_{\omega\in B}$ المجموع فوق نقاط العينة ω التي تنتمي إلى B. وهذه العلاقة تعبر عن حقيقة أن B هي الآن (تحت شرط وقوع B) الحادثة الأكيدة، مما يجعل احتمال أي نقطة عينة لا تنتمي إلى B مساويا للصفر وفقا للدالة الشرطية P_B، ويزيد من احتمال كل نقطة تنتمي إلى B بمقدار يتناسب مع الاحتمال الذي خصتها به الدالة غير الشرطية P. الفكرة التي أوضحناها في سياق المثال (٢ـ٢٥). وهذا يسمح لنا بكتابة :

$$P_B(\{\omega\}) = \begin{cases} KP(\{\omega\}) , & \forall\ \omega \in B, \\ 0 , & \forall\ \omega \notin B. \end{cases}$$

حيث K عدد ثابت موجب. ولكن

$$1 = \sum_{\omega\in B} P_B(\{\omega\}) = K \sum_{\omega\in B} P(\{\omega\}) = KP(B)$$

وبالتالي،

$$K = \frac{1}{P(B)}$$

والعلاقة السابقة تصبح :

$$P_B(\{\omega\}) = \begin{cases} \dfrac{P(\{\omega\})}{P(B)} , & \omega \in B ; \\ 0 , & \omega \in \bar{B} . \end{cases}$$

والآن، من أجل أي حادثة A، لدينا :

$$\begin{aligned} P(A|B) = P_B(A) &= \sum_{\omega\in A} P_B(\{\omega\}) \\ &= \sum_{\omega\in A\cap B} P_B(\{\omega\}) + \sum_{\omega\in A\cap \bar{B}} P_B(\{\omega\}) \end{aligned}$$

$$= \sum_{\omega \in A \cap B} P_B(\{\omega\}) + 0$$

$$= \sum_{\omega \in A \cap B} \frac{P(\{\omega\})}{P(B)} = \frac{1}{P(B)} \sum_{\omega \in A \cap B} P(\{\omega\})$$

$$= \frac{P(A \cap B)}{P(B)}$$

(٢ـ١٤) الاستقلال

لتكن A، B، حادثتين من فضاء عينة S. ولنفرض أننا حسبنا $P(A|B)$ فوجدناه مساويا لـ $P(A)$ ، فماذا نقول عن حالة كهـذه؟ حساباتنا تشير إلى أن وقوع B لم يكن له أثر على احتمال وقـوع A، وقـد ذكرنـا في مطلع الفقرة السـابقة أنـه من الطبيعي وصف الحادثتين بأنهما مستقلتـان احتماليا . وسنكتفـي من الآن فصاعـدا بـالقول إن حـادثتين مستقلتان ، ونقصد بالطبع أن الحادثتين مستقلتان احتماليا .

لنكتب الآن التعبير الرمزي عن استقلال حادثتين A ، B ، أي :

$$P(A \mid B) = P(A)$$

ولنعوض عن $P(A \mid B)$ بما يساويها وفقا لتعريف الاحتمال الشرطي فنجد :

$$\frac{P(AB)}{P(B)} = P(A) \quad , \quad P(B) \neq 0$$

أو

$$P(AB) = P(A)\,P(B)$$

وعلى العكس، لو فرضنا أن $P(B) \neq 0$ ، وأن :

$$P(AB) = P(A)\,P(B)$$

فنجد بقسمة الطرفين على P(B) وتطبيق تعريف الاحتمال الشرطي أن :

$$P(A \mid B) = P(A)$$

أي أن الحادثتين A و B مستقلتان . ومنه نستنتج القاعدة التالية :

الشرط اللازم والكافي لاستقلال حادثتين A و B هو أن يكون

$$P(AB) = P(A)\,P(B)$$

وهذه القاعدة تقول، إذا كنا نعلم أن حادثتين A و B مستقلتان فاحتمال وقوعهما معا هو جداء احتماليهما. وكي نقرر في مسألة استقلال أو عدم استقلال حادثتين A ، B نحسب احتمال وقوع كل منهما، $P(A)$ ، $P(B)$ ، ونحسب احتمال وقوعهما معا، $P(AB)$ ، فإذا وجدنا أن الشرط المذكور أعلاه محقق استنتجنا أنهما مستقلتان، وإذا وجدنا أنه غير محقق استنتجنا أنهما غير مستقلتين. وهذا يدعو إلى تبني هذا الشرط كتعريف لاستقلال حادثتين.

(٢ ـ ١٤ ـ ١) الحادثتان المستقلتان

نقول إن الحادثتين A و B مستقلتان إذا، وفقط إذا، كان:

$$P(AB)=P(A)\,P(B)$$

مثال (٢ ـ ٢٨)

لنعد إلى مثال قذف حجر النرد في الفقرة السابقة حيث وجدنا أن $P(A_3|B) = P(A_3) = 1/3$ ، فالحادثة A_3 مستقلة عن الحادثة B. ونلاحظ تحقق الشرط:

$$P(A_3 B) = P(A_3)\,P(B),$$

$$\frac{1}{3} = \frac{2}{3} \times \frac{1}{2}$$

ولكن A_2 غير مستقلة عن B لأن

$$P(A_2 B) \neq P(A_2)\,P(B)$$

$$\frac{1}{6} \neq \frac{1}{2} \times \frac{1}{2}.$$

وكذلك A_1 غير مستقلة عن B ، (تحقق من ذلك).

مثال (٢ ـ ٢٩)

في صندوق تسع قطع نقود من الأنواع المبينة في الجدول التالي وتحمل التواريخ المبينة لكل نوع.

ربع ريال 1976 ، 1978 ، 1980 ، 1982

نصف ريال 1976 ، 1980 ، 1982

ريـــال 1980 ، 1983 .

سحبنا قطعـة بصورة عشـوائية ، لتكن A حـادثة سحب ربع ريـال ؛ B حادثة سحب نصف ريـال ؛ و C حادثـة سحب قطعة نقـود تحمل التاريخ 1980 ، والمطلوب حساب : $P(C)$ ، $P(A|C)$ ، هل الحادثتان A و C مستقلتان؟ هل الحادثتان B و C مستقلتان؟

الحـــل

لحساب احتمال C نلاحظ أن عدد الحالات الممكنة 9، وعدد الحالات الملائمة 3 ويكون

$$P(C) = \frac{3}{9} = \frac{1}{3}$$

وبصورة مماثلة نجد أن

$$P(AC) = \frac{1}{9}$$

ووفقا لتعريف الاحتمال الشرطي يكون

$$P(A|C) = \frac{P(AC)}{P(C)} = \frac{1/9}{1/3} = \frac{1}{3}$$

وللحكم في استقلال A ، C نحسب $P(A)$ والجداء $P(A) \times P(C)$ ثم نقارنه مـع $P(AC)$. فنجد

$$P(A) = \frac{4}{9}$$

$$P(A) \times P(C) = \frac{4}{9} \times \frac{1}{3} = \frac{4}{27} \neq P(AC) = \frac{1}{9}$$

فالحادثتان A ، C غير مستقلتين .

وللحكم في استقلال B ، C نحسب ، بصورة مماثلة ،

$$P(B) = \frac{3}{9} = \frac{1}{3}$$

$$P(BC) = \frac{1}{9}$$

$$P(B) \times P(C) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9} = P(BC)$$

فالحادثتان B ، C مستقلتان.

مثال (٢ ـ ٣٠)

الحادثتان A و B متنافيتان ، و $P(A) \neq 0$ ، و $P(B) \neq 0$. ادرس استقلال الحادثتين.

الحل

بما أن الحادثتين متنافيتان فإن تقاطعهما خال. (لا يمكن وقوعهما معا) أي $P(AB) = P(\phi) = 0$. ولا يمكن تحقق شرط الاستقلال، لأن أحد الطرفين ($P(AB)$ يساوي الصفر، والطرف الآخر $P(A)\,P(B)$، هو جداء عددين موجبين بالفرض، أي أنه لا يمكن أن يساوي صفرا. وهكذا نستنتج أن الحادثتين المتنافيتين هما على وجه التأكيد، غير مستقلتين. وهذه النتيجة تنسجم تماما مع بداية كلامنا عن الاستقلال، فوقوع أحدهما يجعل احتمال وقوع الأخرى صفرا، وأي تأثير يمكن أن يكون أكبر من ذلك!

(٢ ـ ١٥) قانونان أساسيان في الاحتمال واستخدامهما

(٢ ـ ١٥ ـ ١) قانون الجمع

برهنا في النتيجة (٢ ـ ٨ ـ ٧) أن:

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

وهو ما يسمى بقانون الجمع.

(٢ ـ ١٥ ـ ٢) قانون الجداء

من تعريف الاحتمال الشرطي نجد أن

$$P(AB) = P(A)\ P(B \mid A)$$
$$= P(B)\ P(A \mid B)$$

وهو قانون الجداء.

وتجدر ملاحظة أن قانون الجمع يصبح مسلمة الاحتمال الثالثة عندما تكون الحادثتان A، B، متنافيتين، أي $AB = \phi$.إذ يصبح الحد الثالث $P(AB)$ صفرا. كما تجدر ملاحظة أن قانون الجداء يصبح، في حالة استقلال الحادثتين A، B، الشرط اللازم والكافي لاستقلالهما. إذ يكون عندئذ $P(A \mid B) = P(A)$ و $P(B \mid A) = P(B)$.

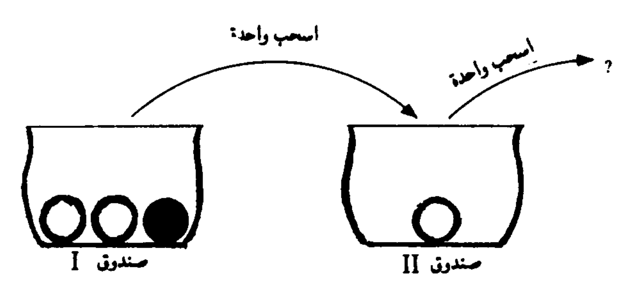
مثال (٢ ـ ٣١)

لنعد الآن إلى المثال (٢ ـ ١٥) احسب باستخدام القواعد والقوانين الأساسية التي تعلمتها، احتمال الحصول على كرة بيضاء من الصندوق الثاني.

الحل

لنرمز بـ A لحادثة الحصول على كرة بيضاء من الصندوق الثاني. فلا يمكن الوصول إلى A إلا بإحدى طريقتين:

أن نسحب «كرة بيضاء من الصندوق I» و «كرة بيضاء من الصندوق II» (ولنرمز لهذه الحادثة بـ B). أو أن نسحب «كرة سوداء من الصندوق I» و «كرة بيضاء من الصندوق II» (ولنرمز لهذه الحادثة بـ C). ونلاحظ أن B و C متنافيتان وأن $A = B \cup C$.
(أي أن A تتحقق بوقوع B أو C).



شكل (٢ ـ ١٢): تمثيل للتجربة في المثال (٢ ـ ٣١)

ومن عبارة B نلاحظ أن $B = B_1 A$ حيث B_1 حادثة سحب كرة بيضاء من الصندوق I. كما نلاحظ من عبارة C أن $C = C_1 A$ حيث C_1 حادثة سحب كرة سوداء من الصندوق I. ويمكننا أن نكتب الآن، اعتمادا على قوانين معروفة،

$$
\begin{aligned}
P(A) &= P(B \cup C) \\
&= P(B) + P(C) \qquad (B \text{ و } C \text{ متنافيتان}) \\
&= P(AB_1) + P(AC_1) \\
&= P(A|B_1)\,P(B_1) + P(A|C_1)\,P(C_1) \\
&= \frac{2}{3} \times \frac{2}{3} + \frac{1}{3} \times \frac{1}{3} = \frac{4}{9} + \frac{1}{9} = \frac{5}{9}
\end{aligned}
$$

وهو الجواب نفسه الذي حصلنا عليه في حل المثال (٢ ــ ١٥) مستخدمين هناك فضاء العينة وتعريف احتمال حادثة.

مثال (٢ ـ ٣٢)

بالعودة إلى المثال (٢ ـ ١٦) حيث اخترنا عشوائيا بذرتين من خمس بذور. ما هو احتمال الحصول على بذرة تنتج زهورا بيضاء وبذرة تنتج زهورا حمراء؟

الحــــل

لنرمز بـ A لحادثة الحصول على بذرة تنتج زهورا بيضاء وبذرة تنتج زهورا حمراء فيمكن الوصول إلى A بإحدى طريقتين، فإما أن نختار بذرة الزهور البيضاء أولا وبذرة الزهور الحمراء ثانيا (ولنرمز لهذا الطريق بـ B)، أو نختار بذرة الزهور الحمراء أولا وبذرة الزهور البيضاء ثانيا (ولنرمز لهذا الطريق بـC).

ومن عبارتي B و C نلاحظ أن $B = B_1 B_2$، $C = C_1 C_2$. حيث ترمز B_1 لحادثة اختيار بذرة الزهور البيضاء أولا وB_2 لحادثة اختيار بذرة الزهور الحمراء ثانيا وترمز C_1 لحادثة اختيار بذرة الزهور الحمراء أولا و C_2 لحادثة اختيار بذرة الزهور البيضاء ثانيا ويكون

$$A = B \cup C = B_1 B_2 \cup C_1 C_2$$

$$P(A) = P(B \cup C) = P(B) + P(C) \quad \text{(} B \text{ و } C \text{ متنافيتان)،}$$

$$= P(B_1 B_2) + P(C_1 C_2)$$

$$= P(B_1)\, P(B_2 | B_1) + P(C_1)\, P(C_2 | C_1)$$

$$= \frac{2}{5} \times \frac{2}{4} + \frac{2}{5} \times \frac{2}{4} = \frac{8}{20} = 0.4$$

حل آخـــــر

باستخدام طرق العدّ، نلاحظ بسهولة أن عدد الحالات الملائمة هو عدد إمكانات اختيار بذرة زهور بيضاء، وبذرة زهور حمراء. ولكن يمكن اختيار بذرة زهور بيضاء بطريقتين مختلفتين وفي كل منها يمكن اختيار بذرة زهور حمراء بطريقتين مختلفتين أيضا، ويكون عدد الحالات الملائمة $4 = 2 \times 2$ وعدد الحالات الممكنة هو عدد طرق اختيار بذرتين من خمس بذور ويساوي $C_2^5 = 10$. والاحتمال المطلوب هو:

$$\frac{4}{10} = 0.4$$

مثال (٢ ـ ٣٣)

احتمال أن يكون باب معين مقفلا هو 1/2. ومفتاح الباب هو بين 12 مفتاحا متوفرة ضمن حزمة واحدة إذا اختار شخص مفتاحين بصورة عشوائية، فما هو احتمال أن يستطيع فتح الباب دون اللجوء إلى مفاتيح أخرى؟

الحـــل

لنرمز بـ A لحادثة «فتح الباب». ولنتساءل ما هي الطريق التي تؤدي إلى A؟ من الواضح أن A تتحقق إذا وفقط إذا كان الباب غير مقفل أو كان الباب مقفلا واخترنا المفتاح الصحيح. لنرمز الآن لحادثة «الباب مقفل» بـ B، ولحادثة «اختيار المفتاح الصحيح» بـ C . فيمكننا كتابة:

$$A = \bar{B} \cup BC$$

ومن الواضح أن B و C مستقلتان، وأن $\bar{B}$ و BC متنافيتان، وبالتالي:

$$P(A) = P(\bar{B} \cup BC) = P(\bar{B}) + P(BC)$$
$$= 1 - P(B) + P(B)\,P(C)$$

ولكن $P(B) = 1/2$، ولحساب احتمال C نقوم بالمحاكمة التالية:

تتحقق C إذا، وفقط إذا، كان أحد المفتاحين اللذين اخترناهما هو المفتاح الصحيح، ويمكن اختيار المفتاح الصحيح بطريقة واحدة، واختيار المفتاح غير الصحيح بـ 11 طريقة، ويكون عدد الحالات الملائمة $1 \times 11 = 11$، وعدد الحالات الممكنة لاختيار مفتاحين هو C_2^{12} وبالتالي:

$$P(C) = \frac{11}{C_2^{12}} = \frac{11 \times 2}{11 \times 12} = \frac{2}{12} = \frac{1}{6}$$

والآن

$$P(A) = 1 - \frac{1}{2} + \frac{1}{2} \times \frac{1}{6} = \frac{1}{2} + \frac{1}{12} = \frac{7}{12}$$

(٢ ـ ١٦) التكرارات المستقلة

إذا كانت الحوادث $A_1, A_2, \ldots, A_n$ مستقلة فيما بينها فيمكن أن نكتب كتعميم لما وجدناه في حالة استقلال حادثتين:

$$P(A_1 . A_2 \ldots A_n) = P(A_1)\,P(A_2) \ldots P(A_n)$$

وتنبغي ملاحظة أن تحقق هذه العلاقة لا يكفي للقول باستقلال هذه الحوادث بعضها عن بعض. إذ يجب تحقق شروط أخرى إضافية سوف لا ندخل هنا في تفاصيلها. ولكن ما قلناه لا يتعدى أنه إذا كانت الحوادث مستقلة فيما بينها، فإن هذه العلاقة تكون صحيحة.

مثال (٢ ـ ٣٤)

قذفنا قطعة نقود ثلاث مرات متتالية. احسب احتمال:

أ ـ الحصول على HHT،

ب ـ الحصول على وجه الـ H مرتين.

الحــــل:

يتضح من طبيعة التجربة أنه لا يمكن أن يكون لنتيجة إحدى القذفات، أي أثر في الاحتمالات الموافقة لنتائج قذفة أخرى. والقذفات الثلاث هي تكرارات مستقلة للتجربة نفسها. وفي كل تكرار نعلم أن $P(H) = P(T) = 1/2$.

أ ـ $P(HHT) = P$ (H في القذفة الأولى و H في القذفة الثانية و T في القذفة الثالثة)

$= P$(H في القذفة الأولى) $\times$ P (H في القذفة الثانية)

$\times$ P(T في القذفة الثالثة)

$$= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

ب ـ حادثة «الحصول على وجه الـ H مرتين ولنرمز لها بـ A يمكن أن تتحقق بثلاثة أشكال مختلفة هي HHT أو HTH أو THH وهكذا نكتب:

$P(A) = P(HHT) + P(HTH) + P(THH)$، (حسب المسلمة الثالثة)

$$= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

كيف علمنا بوجود ثلاثة أشكال مختلفة تحقق المطلوب؟

الجواب: عدد هذه الأشكال هو عدد إمكانات اختيار موقعين من ثلاثة مواقع لنضع فيها H ونترك الباقي لـ T. وهذا العدد كما نعلم من الطرق العد هو $C_2^3 = 3$.

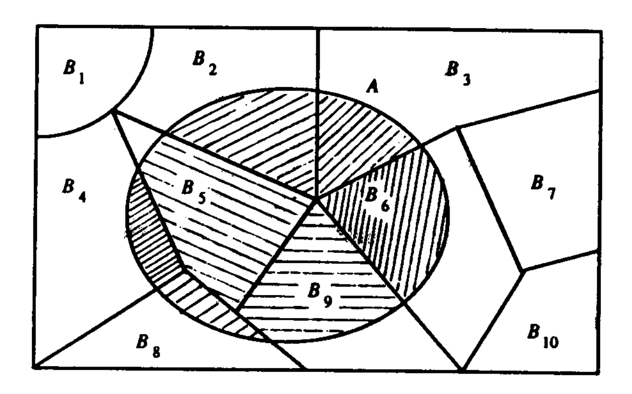
(٢ ـ ١٧) الاحتمال الكلي

لنفرض أن الحوادث غير الخالية $B_1, B_2, ..., B_k$ تشكل تجزئة لفضاء عينة S. أي أنها متنافية ومستنفذة ($B_i \cap B_j = \phi$ لأي $i \neq j$ ؛ و $B_1 \cup B_2, ... \cup B_k = S$) فيمكن

التعبير عندئذ عن أي حادثة A من S بدلالة تقاطعات هذه الحادثة مع كل من حوادث التجزئة. وهذا واضح مما يلي:

$$A = A \cap S = A \cap (B_1 \cup B_2 \cup \ldots \cup B_k)$$
$$= (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k)$$

(انظر الشكل ٢ ـ ١٣)



شكل (٢ ـ ١٣) عشر حوادث B_1 إلى B_{10} تشكل تجزئة لفضاء عينة S .

ووفقا للمسلمة الثالثة نجد:

$$P(A) = P(AB_1) + P(AB_2) + \ldots + P(AB_k)$$

وبتطبيق قانون الجداء على كل حد من حدود الطرف الأيمن نجد:

$$P(A) = P(A \mid B_1)\, P(B_1) + P(A \mid B_2)\, P(B_2) + \ldots + P(A \mid B_k)\, P(B_k)$$

وهو قانون الاحتمال الكلي. ويمكن كتابته باختصار كما يلي:

$$P(A) = \sum_{i=1}^{k} P(A \mid B_i)\; P(B_i)$$

مثال (٢ ـ ٣٥)

مصنع للجوارب يتضمن ثلاث آلات. مساهمة كل منها في الإنتاج الكلي اليومي للمصنع هي، على الترتيب، 30%، 36%، 34%. اخترنا عشوائيا جوربا من الإنتاج

الكلي اليـومي للمصنع . ما هـو احتمال أن يكون معيبـا (فيه عيب صنـاعي) ، علما أن النسبة المئوية للإنتاج المعيب في الآلات الثلاث هي ، على الترتيب ، 1% ، 2%، و2% ؟

الحـــل

لنرمز بـ :

B_1 لحادثة «الجورب من إنتاج الآلة الأولى» ،

B_2 لحادثة «الجورب من إنتاج الآلة الثانية ،

B_3 لحادثة ««الجورب من إنتاج الآلة الثالثة» ،

A لحادثة «الجورب معيب» .

نلاحظ أولا أن B_1 ، B_2 ، B_3 تشكل تجزئة لفضاء العينة S الموافق لتجربة الاختيار العشوائي لجورب من مجمل الانتاج اليومي للمصنع . فأي جـورب نختاره لابد أن يكون من إنتاج الآلـة الأولى ، أو من إنتاج الآلة الثانيـة ، أو من إنتاج الآلة الثـالثة . وبتطبيق قانون الاحتمال الكلي نجد :

$$P(A) = P(A|B_1)\,P(B_1) + P(A|B_2)\,P(B_2) + P(A|B_3)\,P(B_3)$$

ولكن من معطيات المسألة نلاحظ أن :

$$P(B_1) = 0.30\,;\ P(B_2) = 0.36\,;\ P(B_3) = 0.34$$

(لاحظ أن مجموع احتمالات حوادث التجزئة يجب أن يكون مساويا للواحد) .

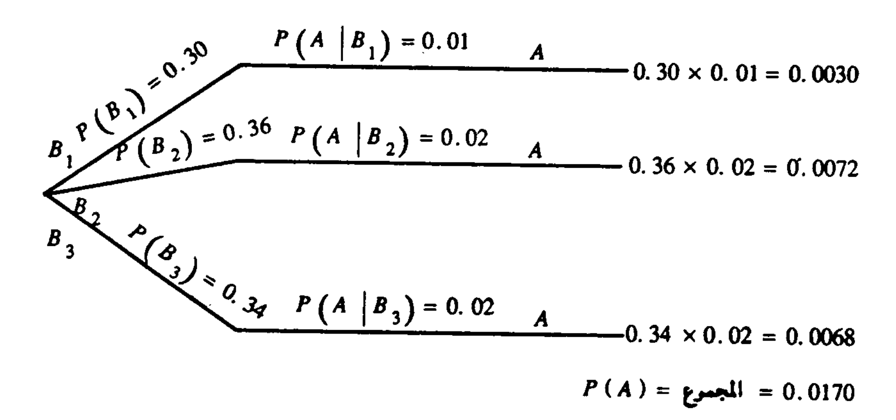
$$P(A|B_1) = 0.01\,;\ P(A|B_2) = 0.02\,;\ P(A|B_3) = 0.02$$

وبالتعويض في علاقة الاحتمال الكلي نجد :

$$P(A) = 0.01 \times 0.30 + 0.02 \times 0.36 + 0.02 \times 0.34 = 0.017$$

ملاحظة

يوضح المخطط في الشكل (٢ ـ ١٤) المسألة في المثال السابق . ويسمى مثل هذا المخطط ، عادة ، مخطط الشجرة .



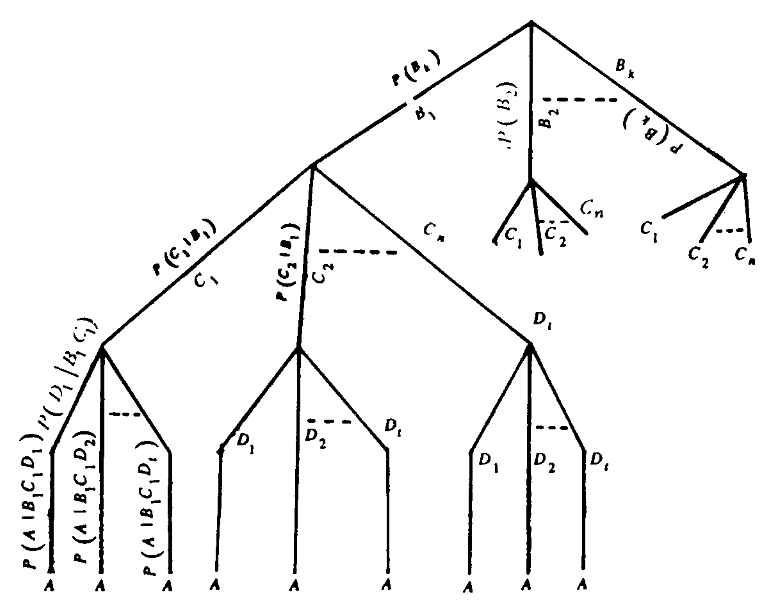
شكل (٢ ـ ١٤) : مخطط الشجرة لحل المثال (٢ ـ ٣٥)

(٢ ـ ١٧ ـ ١) طريقة مخطط الشجرة لحل مسألة احتمالية

يمكن تعميم فكرة مخطط الشجرة التي استعرضناها لحل المثال (٢ ــ ٣٥) إلى مسائل احتمالية تتعدد فيها المسارات المؤدية إلى الحادثة المطلوبة، ويتألف كل مسار من عدة غصون، غصن لكل مرحلة من مراحل التجربة. ويمكن تلخيص الطريقة كما يلي: (انظر الشكل ٢ ـ ١٥ التوضيحي).

نرسم غصون المرحلة الأولى بجميع أشكالها الممكنة ونحسب احتمال كل منها، (مجموع هذه الاحتمالات يجب أن يساوي الواحد). ومن كل غصن من أغصان المرحلة الأولى، نرسم كل ما يمكن أن يتفرع من أغصان المرحلة الثانية، ونحسب لكل غصن منها احتمالها الشرطي في ضوء الغصن الذي سبقه (ومجموع هذه الاحتمالات لفروع كل غصن من أغصان المرحلة الأولى يجب أن يساوي الواحد أيضا). وهكذا. . . حتى نصل إلى المرحلة الأخيرة التي تؤدي إلى الحادثة المطلوبة، A مثلا، وفي هذه المرحلة الأخيرة لا يتفرع من كل غصن من أغصان المرحلة السابقة إلا الغصن (الأغصان) التي تؤدي إلى الحادثة A. ونحسب الاحتمال الشرطي الموافق له (لكل منها) في ضوء جميع الغصون السابقة له (لكل منها) والتي تشكل بدءا من المرحلة الأولى وانتهاء بالمرحلة الأخيرة مسارا مؤديا إلى A.

ونحسب الآن لكل مسار احتمالا، هو جداء الاحتمالات المحسوبة لكل غصن من غصونه. وأخيرا نجمع احتمالات المسارات المختلفة فنحصل على احتمال الحادثة A المطلوبة.



شكل (٢ ـ ١٥): رسم توضيحي لطريقة مخطط الشجرة

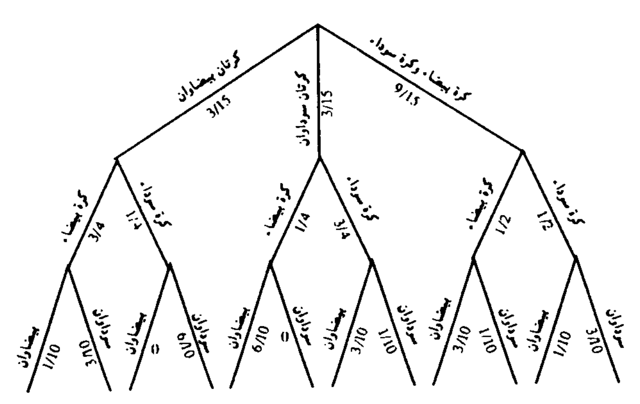
يقتصر المخطط على أربع مراحل تتألف المرحلة الأولى من k غصنا هي $B_1, B_2, ..., B_k$ ويتفرع من كل منها، في المرحلة الثانية n غصنا هي $c_1, c_2, ..., c_n$، ومن كل من الـ n x k غصنا الناتجة يتفرع في المرحلة الثالثة t غصنا هي $D_1, D_2, ..., D_t$. ومن كل من الـ $t \times n \times k$ غصنا الناتجة نأخذ في المرحلة الأخيرة الغصن الذي يؤدي إلى A ولدينا إذا $t \times n \times k$ مسارا وكل مسار مؤلف من أربعة أغصان متتالية، مثلا، المسارات $B_1\ C_1\ D_1\ A$، $B_1\ C_1\ D_2\ A$ . . .، $B_1\ C_1\ D_t\ A$، وهكذا. . .

واحتمال المسار $B_1\ C_1\ D_1\ A$، مثلا، هو:

$$P(B_1\ C_1\ D_1\ A) = P(A |\ B_1\ C_1\ D_1)\ P(D_1 |\ C_1\ B_1)\ P(C_1 |\ B_1)\ P(B_1)$$

مثال (٢ ـ ٣٦)

لدينا في الصندوق I ثلاث كرات بيض وثلاث كرات سود. وفي الصندوق II لدينا كرة بيضاء وكرة سوداء. اخترنا عشوائيا كرتين من الصندوق I ثم خلطناهما جيدا مع كرات الصندوق II. واخترنا من الخليط، عشوائيا، كرة واحدة خلطناها جيدا مع الكرات المتبقية في الصندوق I، ثم اخترنا منه كرتين. احسب احتمال أن تكونا من لون واحد؟



الاحتمال المطلوب:

$$\frac{3}{15}\times\frac{3}{4}\times\frac{1}{10}+\frac{3}{15}\times\frac{3}{4}\times\frac{3}{10}+\frac{3}{15}\times\frac{1}{4}\times 0+\frac{3}{15}\times\frac{1}{4}\times\frac{6}{10}+\frac{3}{15}\times\frac{1}{4}$$
$$\times\frac{6}{10}+\frac{3}{15}\times\frac{1}{4}\times 0+\frac{3}{15}\times\frac{3}{4}\times\frac{3}{10}+\frac{3}{15}\times\frac{3}{4}\times\frac{1}{10}+\frac{9}{15}\times\frac{1}{2}\times\frac{3}{10}$$
$$+\frac{9}{15}\times\frac{1}{2}\times\frac{1}{10}+\frac{9}{15}\times\frac{1}{2}\times\frac{1}{10}+\frac{9}{15}\times\frac{1}{2}\times\frac{3}{10}=$$
$$\frac{9+27+18+18+27+9+54+18+18+54}{600}=\frac{252}{600}=\frac{42}{100}=0.42$$

(٢ ـ ١٨) قانون بايز (Bayes)

لنفرض في المثال (٢ ــ ٣٥) أننا اخترنا جوربا، بصورة عشوائية، فوجدناه معيبا. ونريد حساب احتمال أن يكون هذا الجورب من إنتاج الآلة الأولى. أي أننا نريد

معرفة الاحتمال الشرطي $P(B_1 | A)$. ونلاحظ أنه يمكـن النظر إلى التجزئة B_1, B_2, B_3 في المثـال (٢ ـ ٣٥)، كـأسباب، وأن النتيجـة التي تهمنـا هي مـا إذا كان الجورب الـذي نختاره معيبا. والاحتمال المطلوب $P(B_1 | A)$ هـو إذا احتمال السبب B_1 علما أن النتيجة كانت A . أو بصياغة أكثـر تعبيرا احتمال أن تكون A (التي وقعت) نتيجـة للسبب B_1 دون غيره من الأسباب. ولذلك يسمى مثل هـذا الاحتمال، أحيانا، الاحتمال السببي. ولدينا من قانون الاحتمال الشرطي.

$$P(B_1 | A) = \frac{P(B_1 A)}{P(A)} = \frac{P(A | B_1)\ P(B_1)}{P(A)}$$

ومن قانون الاحتمال الكلي يمكن تعويض $P(A)$ بما تساويه لنجد أخيرا:

$$P(B_1 | A) = \frac{P(A | B_1)\ P(B_1)}{P(A\ B_1)\ P(B_1) + P(A | B_2)\ P(B_2) + P(A | B_3)\ P(B_3)}$$

وهو قـانون بايـز في حالة وجود ثـلاثة أسباب، أي وجـود تجزئة لـ S تقطعه إلى ثـلاثة أجزاء.

وبالتعويض من المثال (٢ ـ ٣٥) نجد:

$$P(B_1 | A) = \frac{0.01 \times 0.30}{0.01 \times 0.30 + 0.02 \times 0.36 + 0.02 \times 0.34}$$

$$= \frac{0.003}{0.017} = \frac{3}{17}\ .$$

وبصورة عامة، إذا فرضنا k من الأسبـاب، أي تجزئة $B_1, B_2, \ldots, B_k$. وكان المطلوب حساب $P(B_j | A)$ أي احتمال أن الحادثة A التي وقعت كانـت نتيجة للسبب B_j ، دون غيره من الأسباب، نكتب من قانون الاحتمال الشرطي:

$$P(B_j | A) = \frac{P(B_j\ A)}{P(A)} = \frac{P(A | B_j)\ P(B_j)}{P(A)}$$

وبتعويض $P(A)$ في المقام بما يساويها، استنادا إلى قانون الاحتمال الكلي، نجد قانون بايز بصورته العامة:

$$P(B_j | A) = \frac{P(A | B_j)\ P(B_j)}{\sum_{i=1}^{k} P(A | B_i)\ P(B_i)} \quad ; \quad j = 1, 2, \ldots, k.$$

مثال (٢ ـ ٣٧)

في مجتمع من البالغين تبلغ نسبة الإصابة بمرض السكري 8%. واحتمال أن يقرر طبيب معين إصابة شخص بهذا المرض، علما أنه مريض بالفعل، هو 0.95 واحتمال أن يقرر إصابته علما أنه غير مصاب هو 0.02 . ما هو احتمال أن يكون شخص بالغ مريضا بالسكري علما أن الطبيب أنبأه بذلك؟

الحل

نتعرف أولا على حوادث التجزئة، وهي ما سميناه بالأسباب. ومن العلامات المميزة لحوادث التجزئة أن مجموع احتمالاتها يجب أن يكون الواحد. ومن الواضح أنها هنا الاصابة أو عدم الاصابة بالسكري.

لتكن B حادثة الإصابة بمرض السكري، ونعلم من معطيات المسألة أن $P(B) = 0.08$. ولتكن B حادثة عدم الاصابة بمرض السكري، ومن الواضح أن $P(B) = 0.92$. لتكن A حادثة أن الطبيب شخّص الإصابة بالمرض. فلدينا من نص المسألة أن $P(A | B) = 0.95$ ، و $P(A | B') = 0.02$ والمطلوب هو حساب $P(B | A)$ ووفقا لقانون بايز لدينا:

$$P(B | A) = \frac{P(A | B)\ P(B)}{P(A | B)\ P(B) + P(A | B')\ P(B')}$$

$$= \frac{0.95 \times 0.08}{0.95 \times 0.08 + 0.02 \times 0.92} = \frac{0.076}{0.0944} = 0.81$$

تمـــارين (٢ ـ ٥)

١) إذا كانت H حادثة أن يحصل خالـد على تقدير ممتاز، و G حادثة أن يكون متفوقا في الرياضة. عبر بكلمات عما تعنيه الرموز التالية :

أ ـ $P(G|H)$، ب ـ $P(G|H)$، جـ ـ $P(H|G')$ ،

د ـ $P(H'|G)$ ، هـ ـ $P(H'|G')$ ، و ـ $P(G'|H')$.

٢) إذا رمزنا بـ A لحادثة أن يكون شخص مصابا بعمـى الألوان، ورمزنا بـ C لحادثة أن يكون تحت العاشرة من العمر. عبر عن الاحتمالات التالية رمزيا :

ا ـ احتمال أن الشخص تحت العاشرة ومصاب،

ب ـ احتمال أن شخصا تحت العاشرة مصاب،

جـ ـ احتمال أن عمر شخص مصاب عشرة أو أكثر،

د ـ احتمال أن شخصا عمره عشرة أو أكثر غير مصاب بعمى الألوان.

٣) تقـدم سـتون شخصا لـوظيفـة. عنـد تصنيفهم وفقا للشهـادة والخبرة حصلنـا على الجدول التالي :

	يحمل شهادة جامعية	لا يحمل شهادة جامعية
له خبرة سابقة	12	6
بدون خبرة سابقة	24	18

اخترنـا أحد المتقـدمين بصورة عشـوائية. ولنـرمز بــ G لحادثـة أنـه يحمل شهـادة جامعية، وبـ T لحادثة أن له خبرة سابقة.

ا ـ احسب الاحتمالات التالية من الجدول مباشرة :

$P(G)$، $P(T)$، $P(TG)$، $P(G'T)$، $P(T|G)$، $P(G|T)$

ب ـ تحقق أن

$$P(G' | T') = \frac{P(G'\ T')}{P(T')} \qquad P(T\ \ G) = \frac{P(TG)}{P(G)}$$

٤) كجزء من الحملة الدعائية تقدم شركة للصناعات الغذائية جائزة مقدارها خمسون ألف ريالا لواحد ممن يرسلون أسماءهم مكتوبة على طلب اشتراك في المسابقة. ووفقا لرغبة المشترك، يمكنه أيضا أن يرسل مع الطلب، الجزء العلوي من علبة تغليف لأحد منتجات هذه الشركة. وقد تبين من فرز وتصنيف 000 60 طلب اشتراك ما يلي:

بدون الجزء العلوي من علبة تغليف	مع الجزء العلوي من علبة تغليف	
11000	32000	سعودي
9000	8000	مقيم

إذا اختير رابح الجائزة بالقرعة، وكانت C حادثة أن يكون الفائز سعوديا، و B حادثة أن الفائز ممن أرسلوا الجزء العلوي من علبة تغليف. احسب كلا من الاحتمالات التالية:

أ- $P(C)$ ، $P(C')$ ، $P(B)$ ، $P(B')$ ، $P(CB)$ ، $P(C'\ B')$
، $P(C|B)$ ، $P(B|C)$ ، $P(C'|B')$ ، $P(B'|C')$.

ب ـ استخدم النتائج في أ للتحقق مما يلي:

$$P(C|B) = \frac{P(CB)}{P(B)}\ ,\qquad P(C'\ |B') = \frac{P(B'\ C')}{P(B')}$$

$$P(B'\ |C') = \frac{P(B'\ C')}{P(C')}\qquad P(B|C) = \frac{P(B\ C)}{P(C)}$$

٥) لنفرض، في التمرين السابق، أنه أعيدت ترتيبات اختيار الفائز بحيث تتضاعف فرصة من يرسل الجزء العلوي من علبة تغليف. أعط تصورا للترتيب الجديد، وأعد كافة الحسابات المطلوبة في ذلك التمرين.

٦) في التمرين ٩ من مجموعة التمارين (٢ ـ ٢)، احسب :
أ ــ احتمال أن المشترك سوف لا يحصل على جائزة التجويد علما أنه حصل على جائزة التفسير.
ب ــ احتمال أن المشترك سوف يحصل على جائزة التفسير علما أنه لم يحصل على جائزة التجويد.

٧) لدى مدير مركز أبحاث المعلومات التالية : احتمال أن يتم استلام تجهيزات، يحتاجها مشروع معين، في وقتها هو 0.8. واحتمال أن يتم تسليم التجهيزات في وقتها وإتمام المشروع في وقته المحدد هو 0.45.
أ ـ احسب احتمال إتمام المشروع في وقته علما أن التجهيزات قد سُلمت في وقتها.
ب ـ إذا كان احتمال أن يتم المشروع في وقته هو 0.5، وعلمت أن التجهيزات سوف لا تتيسر في وقتها، فكم سيصبح هذا الاحتمال؟

٨) تتولى مراكز التأهيل الطبي في المملكة مهمة تأهيل المرضى المعاقين جسميا. وفيما يلي جدول يبين الحالات الجديدة التي تم تأهيلها لعام ١٤٠٦هـ في كل من مركزي مكة المكرمة والرياض: *

نوع الحالة / المركز	شلل أطفال	بتر أطراف	تشوهات	شلل إربي	حالات متنوعة	المجموع
مركز مكة المكرمة	321	179	193	38	814	1545
مركز الرياض	485	243	680	42	540	1990
المجموع	806	422	873	80	1354	3535

* مأخوذ من التقرير الصحي السنوي لعام ١٤٠٦هـ الصادر عن وزارة الصحة في المملكة العربية السعودية.

إذا اخترنا إحدى الحالات عشوائيا فاحسب احتمال أن تكون :

ا ـ حالة شلل أطفال، ب ـ من مركز مكة المكرمة، جـ ـ من مركز الرياض علما أنها حالة بتر أطراف، د ـ حالة تشوه علما أنها من مركز مكة المكرمة، هـ ـ حالة شلل إربي أو شلل أطفال، و ـ حالة شلل إربي أو شلل أطفال علما أنها من مركز الرياض.

٩) فيما يلي جدول يبين عدد الحجاج وعدد حالات ضربة الحرارة في مكة والمشاعر حسب الجنسية وذلك لعام ١٤٠٦هـ:

الجنسية	عدد الحالات	عدد الحجاج	الجنسية	عدد الحالات	عدد الحجاج	الجنسبة	عدد الحالات	عدد الحجاج
مصري	84	98606	هندي	26	39344	عراقي	10	14551
مغربي	72	22912	سوري	22	15803	يماني	8	43512
تركي	67	54624	سعودي	14	239207	ليبي	8	14509
باكستاني	43	92305	تونسي	10	6887	بنجلاديشي	7	13631
اندونيسي	35	59172	أفغاني	12	4603	لبناني	9	4298
جزائري	34	28093	أردني	10	17165	جنوب افريقيا	5	2498
نيجيري	29	29899	إيراني	10	152149	أخرى	81	103212

ا ـ إذا اخترنا أحد الحجاج عشوائيا فما احتمال أن يكون ممن أصيبوا بضربة الحرارة.*

ب ـ إذا اخترنا حاجا بصورة عشوائية فوجدناه سعوديا، ما احتمال ألا يكون قد أصيب بضربة الحرارة.

جـ ـ إذا اخترنا حاجا بصورة عشوائية فوجدناه ممن أصيبوا بضربة الحرارة، ما هو احتمال أن يكون من إحدى البلاد المذكورة تفصيلا في الجدول ومطلة على البحر الأبيض المتوسط.

* مأخوذ من التقرير الصحي السنوي لعام ١٤٠٦هـ الصادر عن وزارة الصحة في المملكة العربية السعودية.

١٠) أظهر تصنيف لطلبة إحدى الكليات أن 40% منهم من أهالي الرياض، و 80% منهم يتناولون وجبة الغداء بانتظام في مطعم الجامعة، و 30% منهم من أهالي الرياض ويتناولون وجبة الغداء بانتظام في مطعم الجامعة. **

أ - ما هي النسبة المئوية للطلبة من غير أهالي الرياض ولا يتناولون وجبة الغداء بانتظام في مطعم الجامعة.

ب - من بين الطلبة من أهالي الرياض ما هي نسبة الطلاب الذين يتناولون وجبة الغداء بانتظام في مطعم الجامعة؟

جـ - من بين الطلاب الذين لا يتناولون وجبة الغداء بانتظام في مطعم الجامعة ما هي نسبة الطلاب من أهالي الرياض؟

١١) ينتمي ستون بالمائة من الطلبة المسجلين في مقرر الاحصاء 101 إلى كلية العلوم، وينتمي الباقون إلى كلية الحاسب الآلي. وكانت نسبة النجاح في هذا المقرر هي 70% بالنسبة إلى طلاب كلية العلوم، بينما ترتفع هذه النسبة إلى 90% بين طلاب الحاسب الآلي:

أ - اخترنا طالبا بصورة عشوائية، فما احتمال أن يكون ناجحا؟

ب - إذا علمت أن الطالب الذي اخترناه كان من الناجحين، فما احتمال أنه من طلاب كلية الحاسب الآلي؟

١٢) أي الأزواج التالية من الحوادث مستقل وأيها غير مستقل؟

ا - أن يكون سائق سيارة مخمورا، وأن يرتكب حادث اصطدام،

ب - الحصول على ثلاث ثم ثلاث في قذفتين متتاليتين لحجر نرد،

جـ - أن يكون شخص مدير مصرف، وأن يكون أسود الشعر،

د - حصول بنشر لسيارتك، وتأخرك عن موعد عملك،

هـ - أن يكون شخص من مواليد يوليو (تموز) وأن تكون قدماه مسطحتين،

و - أن يكون لديك رخصة قيادة، وأن تمتلك سيارة،

ز - أن تكون ممن يعيشون في الرياض، ومن هواة جمع الطوابع،

ح - أي حادثتين متنافيتين وغير مستحيلتين.

** النسب المعطاة افتراضية وليست حقيقية.

١٣) في المثال (٢ ـ ٤)، افترض أن الاحتمالات المخصصة لنقاط العينة كانت ما يلي:

(1,1)	(1,0)	(1,-1)	(0,-1)	(0,0)	(0,-1)	(-1,1)	(-1,0)	(-1,-1)	نقطة العينة
0.16	0.08	0.16	0.08	0.04	0.08	0.16	0.08	0.16	الاحتمـــال

لتكن N حادثة أن الشخص الأول على الحياد، و S حادثة أن الشخص الثاني ضد القضية.

ا ـ احسب $P(N)$، $P(S)$، $P(NS)$، $P(N \mid S)$،

ب ـ تحقق أن الحادثة N مستقلة عن الحادثة S،

جـ ـ تحقق أن الحادثة N مستقلة عن الحادثة S'،

د ـ تحقق أن الحادثة S مستقلة عن الحادثة N'،

١٤) في التمرين ٦ من مجموعة التمارين (٢ ـ ٣)، هل الحادثتان A و T مستقلتان؟

١٥) يحتفظ مستشفى بسيارتي إسعاف احتياطا للطوارىء. ونظرا لتوقيت الطلب أو لإمكانية وجود عطل ميكانيكي، فإن احتمال توفر سيارة إسعاف معينة عند الحاجة إليها هو 0.9. وتوفر إحدى السيارتين مستقل عن توفر الأخرى. والمطلوب:

ا ـ ما احتمال ألا تتوفر أي منهما؟

ب ـ إذا احتجنا لسيارة إسعاف في حالة طارئة فما احتمال تلبية الطلب؟

١٦) الحادثتان A، B مستقلتان. و $P(A) = 0.3$، $P(B) = 0.4$، احسب:

ا ـ احتمال وقوعهما معا،

ب ـ احتمال وقوع واحدة منهما على الأقل،

جـ ـ احتمال وقوع واحدة منهما بالضبط،

د ـ احتمال عدم وقوع أي منهما.

١٧) إذا كان احتمال مولود ذكر يساوي $1/2$. وكان الجنس مستقلا من طفل إلى آخر، فما احتمال أن نجد في أسرة تتضمن أربعة أطفال:

ا - الأطفال الأربعة ذكور؟

ب - أحدهم على الأقل ذكر؟

جـ - عدد الذكور يساوي عدد الإناث؟

١٨) كم مرة يجب قذف قطعة نقود حتى يكون احتمال ملاحظة وجه الـ T مرة واحدة على الأقل أكبر من 0.9؟

١٩) خمس قطع من الورق كُتبت عليها الحروف a، b، c، d، e، حرف على كل قطعة. سحبنا ثلاث قطع عشوائيا. لتكن A حادثة الحصول على الحرف a ولتكن B حادثة الحصول على الحرف b، ولتكن C حادثة عدم الحصول على الحرف d في المجموعة التي اخترناها. احسب:

ا - $P(A)$، $P(B)$، $P(C)$،

ب - هل A و C مستقلتان؟

جـ - احسب $P(B \mid A)$، $P(C' \mid B)$، $P(A \cup C)$

٢٠) في ناد يتضمن ستة أطفال، من بينهم أحمد وخالد. اخترنا بالقرعة لجنة من ثلاثة.

ا - ما هو احتمال أن تتضمن اللجنة أحمدا ولا تتضمن خالدا؟

ب - إذا علمت أن اللجنة تتضمن أحمدا فما هو الاحتمال الشرطي أنها تتضمن خالدا أيضا؟

٢١) أنتجت آلة صناعية 20 قطعة، فوجد أن 12 منها موافقة للطول المطلوب و 5 قطع أكبر من الطول المطلوب، و3 قطع أصغر من الطول المطلوب. سحبت قطعة من هذا الانتاج عشوائيا. احسب احتمالات الحوادث:

ا - القطعة المسحوبة موافقة للطول المطلوب،

ب - القطعة المسحوبة غير موافقة للطول المطلوب،

جـ - القطعة المسحوبة أكبر من الطول المطلوب علما أنها غير موافقة للطول المطلوب.

٢٢) في التمرين السابق، إذا سحبنا قطعتين بدون إعادة، فاحسب احتمالات الحوادث:

ا - القطعتان المسحوبتان موافقتان للطول.

ب - القطعتان المسحوبتان غير موافقتين للطول.

جـ - القطعتان المسحوبتان أكبر من الطول المطلوب.

د - القطعة الأولى موافقة للطول المطلوب والثانية أكبر منه.

هـ - واحدة أكبر من الطول المطلوب، والأخرى أصغر من الطول المطلوب.

٢٣) في التمرين السابق احسب الاحتمالات المطلوبة إذا كان السحب يجري مع الإعادة.

٢٤) عينة تتضمن 24 صماما منها 5 تالفة. سُحبت بدون إعادة عينة من 4 صمامات احسب احتمال:

ا - ألا تتضمن العينة صمامات تالفة،

ب - أن تكون العينة كلها تالفة،

جـ - أن يكون نصف العينة تالفا،

د - أن تتضمن العينة قطعة واحدة تالفة.

٢٥) حل التمرين السابق إذا كان السحب مع الإعادة.

٢٦) نعلم أن احتمال وقوع أي عدد من الحوادث، المستقلة فيما بينها، يساوي جداء احتمالاتها. استخدم هذه القاعدة لحساب احتمال:

ا - الحصول على وجه الـT ثماني مرات متتالية عند قذف قطعة نقود متزنة ثماني مرات.

ب - الحصول على وجه الـH في القذفات الأربع الأولى ثم الحصول على وجه الـT في القذفات الأربعة التالية.

جـ - الحصول على 5 أو 6 ثلاث مرات عند قذف حجر نرد متوازن أربع مرات.

د - أن يصيب رام الهدف خمس مرات متتالية علما أن احتمال إصابته للهدف في كل مرة 0.9، وأنه يمكن افتراض الاستقلال بين رمية وأخرى.

٢٧) حزمتان من البطاريات تحوي كل منهما ست بطاريات . وفي كل منهما بطاريتان لا تعملان . إذا اخترنا بطاريتين من كل حزمة فما احتمال أن تكون البطاريات الأربعة عاملة؟

٢٨) إذا علمت أن الصندوق I فيه ثلاث كرات بيض وخمس كرات سود ، وفي الصندوق II خمس كرات بيض وثلاث كرات سود . وسحبنا مع الاعادة كرتين من الصندوق I ، وبدون إعادة كرتين من الصندوق II ، فما هو احتمال الحصول على :

ا ـ 4 كرات بيض

ب ـ كرتين بيضاوين

جـ ـ كرة سوداء واحدة على الأقل .

٢٩) بالاشارة إلى التمرين ٢٥ . لنفرض أننا اخترنا بصورة عشوائية بطاريتين من الحزمة الأولى وخلطناهما مع بطاريات الحزمة الثانية ، ثم أخذنا بصورة عشوائية اثنتين من البطاريات الثماني في الحزمة الثانية ، فما هو احتمال أن تكونا عاملتين؟

٣٠) يتضمن صندوق ثلاث كرات حمراء وأربع كرات بيضاء وخمس كرات زرقاء ، ويتضمن صندوق آخر كرة حمراء وست كرات بيضاء وثلاث كرات زرقاء . سحبنا عشوائيا كرة من كل صندوق . احسب احتمالات الحوادث :

ا ـ الكرتان من اللون نفسه ،

ب ـ واحدة حمراء وواحدة بيضاء ،

جـ ـ واحدة حمراء على الأقل ،

د ـ كلاهما ليست زرقاء .

٣١) يقوم مصنع بتنفيذ دورات تدريبية لمعظم عماله الجدد . ونعلم من سجلات المصنع أن 35% من بين العمال الجدد الذين لم يتلقوا الدورة التدريبية يحسنون أداء عملهم ، بينما ترتفع هذه النسبة إلى 86% بين العمال الجدد الذين تلقوا الدورة التدريبية . إذا علمت أن 80% من العمال الجدد في المصنع تلقوا دورة تدريبية . فما احتمال أن عاملا اخترناه عشوائيا من بين العمال الجدد سيُحسن أداء عمله؟

٣٢) يستأجر فندوق سيارات لنزلائه من 3وكالات X، Y، Z، وذلك وفق النسب التالية: 20% من X، و40% من Y، و40% من Z . إذا كان 14% من سيارات X، و4% من سيارات Y، و8% من سيارات Z تفتقر إلى مذياع، فما احتمال أن سيارة استؤجرت لأحد النزلاء تفتقر إلى مذياع؟

٣٣) احتمال أن يشترك مقاول A في مناقصة لبناء دار جديدة لبلدية إحدى المدن هو 1/2. اشترك المقاول B في المناقصة، واحتمال أن يفوز بالعقد هو 2/3 في غياب المقاول A، ويصبح 1/5 فقط عند اشتراك المقاول A في المناقصة. إذا علمت أن المقاول B قد فاز بالعقد فما احتمال أن المقاول A لم يشترك في المناقصة؟

٣٤) في مكتب للبريد ثلاثة أقسام هي S, Q, R تقوم بتصنيف وتوزيع الخطابات. ونعلم من السجلات السابقة للمكتب أن S يرتكب خطأ واحدا في كل مائة خطاب، وأن B يرتكب خمسة أخطاء في كل مائة خطاب، أما R فيرتكب ثلاثة أخطاء في كل مائة خطاب. كما نعلم أن العمل موزع بين الأقسام الثلاثة بحيث يقوم S بتصنيف وتوزيع 30% من الخطابات بينما يقوم Q بتصنيف وتوزيع 40% منها، ويتولى R الباقي. في حالة حدوث خطأ، ما هو احتمال أن يكون Q مسؤولا عنه؟

٣٥) تتوزع أبقار مزرعة بين أنواع ثلاث A، B، C، وفق النسب التالية، 25% من النوع A، 35% من النوع B، و40% من النوع C . ونعلم أن 2/3 الأبقار من النوع A، و1/2 الأبقار من النوع B، و 1/4 الأبقار من النوع C، يعطي أكثر من 10 كغ حليب يوميا.

ا - اختيرت بقرة من أبقار المزرعة عشوائيا فوجد أنها تعطي أكثر من 10 كغ حليب يوميا. ما احتمال أن تكون من النوع A؟

ب - اختيرت بقرة عشوائيا فتبين أنها تعطي ما لا يزيد عن 10 كغ حليب يوميا، ما احتمال أن تكون من النوع B؟

٣٦)* توضح سجلات الشرطة أن 30% من حوادث الانفجارات تقع بسبب انقطاع مفاجىء في التيار الكهربائي، وأن 15% منها يقع بسبب ضعف أحد الأجهزة

* النسب المعطاة افتراضية

الكهربائية، وأن 50% يقع بسبب اشتعال أحد الأسلاك، وأن 5% يقع بفعل فاعل. ونعلم من تقديرات الخبراء أن احتمال وقوع الانفجار عند توافر أحد الأسباب السابقة هو، على الترتيب، 0.25، 0.20، 0.40، 0.75. إذا حصل انفجار فكيف نستخدم قانون بايز لتحديد السبب الأكثر شبهة؟

٣٧) يخطط صديقك لقضاء عطلة الأسبوع في إحدى المناطق السياحية أ أو ب أو جـ ويأخذ قراره بالاختيار كما يلي: يقذف حجر نرد فإذا حصل على عدد زوجي يزور المنطقة أ، وإذا حصل على عدد فردي يقذف قطعة نقود، ويزور المنطقة ب إذا حصل على H والمنطقة جـ إذا حصل على T. ونعلم أن احتمال هطول المطر في كل من المناطق الثلاث هو، على الترتيب، 0.3، 0.4، و0.2. عندما عاد صديقك وجدت الوحل على عجلات سيارته فما هو احتمال أنه زار المنطقة أ؟

حوار مع ملحد من منظور إحصائي

المؤمن: أنت تعتقد أن مختلف الظواهر في أنفسنا وفي هذا الكون من حولنا هي بفعل المصادفة البحتة.

الملحد: نعم.

المؤمن: هل يمكن لظاهرة واحدة من الظواهر أن تكون لغير المصادفة بل بفعل خالق مدبر.

الملحد: بالطبع لا، إذ لو اعتقدت بإمكانية ذلك لانحسب إيماني هذا على جميع الظواهر بلا استثناء. وليس هناك ما يسوغ إمكانية وجود جزئي لمدبر يتناول ظاهرة أو ظواهر معينة ويعجز عن تدبير وتصريف غيرها أو ينصرف عنها.

المؤمن: حسناً. لو أمعنا النظر لوجدنا العديد من الظواهر المستقلة بعضها عن بعض فما هو التأثير المتبادل. مثلاً، بين قدرتك على السمع أو النطق وبين النظام العجيب الذي تسير وفقاً له حياة جماعة من النمل؟ وما هي العلاقة بين النظام المدهش لمملكة النحل وبين مراحل تطور الجنين البشري في رحم الأم؟ وما هي العلاقة بين سرعة دوران الأرض حول نفسها وقدرة الخفافيش على أن تبلغ أهدافها في الظلام الدامس؟ في الحقيقة يمكن أن نستعرض عدداً هائلاً من الظواهر المستقلة في كوكبنا الأرضي وحده، الذي لا يشكل إلا ذرة لا متناهية في الصغر من الكون الفسيح بما يحويه من بلايين المجرات.

الملحد: لا اعتراض لي على ما تقول ولكن ما هو قصدك من ذلك.

المؤمن: لابد أنك سمعت بنظرية تسمى نظرية الاحتمالات، وهي نظرية تنتمي إلى ميدان الرياضيات البحتة. دعنا نسجل n من الظواهر المستقلة ثم نُقيم عليها نموذجاً احتمالياً هو نموذج بيرنوللي. وسأقيم هذا النموذج متحيزاً لصالحك وبالقدر الذي ترغبه.

كل ظاهرة من هذه الظواهر إما أن تكون بفعل المصادفة البحتة كما تقول أو لا تكون. لنفترض أنها بفعل المصادفة البحتة باحتمال عال جداً هو $(1-\varepsilon)$ حيث ε صغير جداً. فهذا النموذج، المنحاز بشدة لصالحك، سيخصص لكل من نقاط فضاءالعينة، وعددها 2^n، احتمالاً. والنقطة الوحيدة التي تخدم أغراضك هي النقطة التي تمثل الحادثة الابتدائية التالية:

جميع هذه الظواهر بدون استثناء هي بفعل المصادفة. والاحتمال المخصص لهذه النقطة. أي احتمال أن يكون هذا صحيحاً هو $(1-\varepsilon)^n$ كما هو معروف جيداً في نظرية الاحتمالات ولا يجادل في هذا اثنان، أما بقية نقاط العينة وعددها 2^n-1 فهي تخدم هدفي. وهي تمثل في جملتها حادثة أنه يوجد على الأقل ظاهرة واحدة من بين هذه الظواهر الـ n ليست بفعل المصادفة، وإنما من تدبير خالق واحد أحد. واحتمال هذه الحادثة هو $1-(1-\varepsilon)^n$ ومن الواضح أن احتمال أن تكون محاكمتك صحيحة وهي $(1-\varepsilon)^n$ يتناهى إلى الصفر مع زيادة n. فيما يتناهى $(1-(1-\varepsilon)^n)$ إلى الواحد، وهو احتمال أن تكون محاكمتي صحيحة. وإليك الآن بعض الحسابات التي توضح ذلك:

$1-\varepsilon$	n	$(1-\varepsilon)^n$
.9	35	.01
0.99	688	.001
.999	9206	.0001
.9999	115124	.00001

فهل هناك أيها الظالم أثر من الحكمة أو المنطق السليم في اتباع محاكمة ينتهي احتمالها إلى الصفر، والإعراض عن محاكمة تنتهي احتمالها إلى الواحد؟

﴿وَلَوْ شَاءَ رَبُّكَ لَآمَنَ مَن فِي الْأَرْضِ كُلُّهُمْ جَمِيعًا أَفَأَنتَ تُكْرِهُ النَّاسَ حَتَّىٰ يَكُونُوا مُؤْمِنِينَ﴾ [يونس: ٩٩]. **﴿وَتَرَى الشَّمْسَ إِذَا طَلَعَت تَّزَاوَرُ عَن كَهْفِهِمْ ذَاتَ الْيَمِينِ وَإِذَا غَرَبَت تَّقْرِضُهُمْ ذَاتَ الشِّمَالِ وَهُمْ فِي فَجْوَةٍ مِّنْهُ ذَٰلِكَ مِنْ آيَاتِ اللَّهِ مَن يَهْدِ اللَّهُ فَهُوَ الْمُهْتَدِ وَمَن يُضْلِلْ فَلَن تَجِدَ لَهُ وَلِيًّا مُّرْشِدًا﴾** [الكهف: ١٧].

الفصل الثالث

المتغير العشوائي والتوزيع الاحتمالي

(٣ ـ ١) مقدمة

رأينا أن التجربة هي أي عملية تـؤدي إلى قياس أو ملاحظة. وعدد الخريجين من جامعة الملك سعود مثلا هو قياس كمـي؛ أو ملاحظة كمية، وأن تفوز الفرس «روعة» في سباق نـادي الفروسية القادم أولا تفـوز ملاحظة وصفية أو كيفيـة. ويمكننا دائما رد المعلومـات الكيفية إلى معلـومات رقمية بتخصيص عـدد لكل نتيجة وصفيـة وفق نظام متفق عليه سلفا، فنسجـل، مثلا، الرقم 1 إذا ربحت «روعة» السبـاق والرقم 0 إذا لم تربحه. وإذا رمزنا لعدد الخريجين بـ X، ولنتيجة «روعة» في السبـاق بـ Y، فمع نهاية كل عام دراسي سنحصل على قيمة للمتغير X، ومع ختام كل سباق تشارك فيه «روعة» سنحصل على قيمـة لـ Y. ومن الطبيعي أن نقـول عن متغير مثل X أو Y إنه متغير عشوائي، لأن القيم التي يفترضها كل منهما مرتبطة بتجارب عشوائية.

مثال (٣ ـ ١)

لتكن التجربـة اختيارا عشوائيـا لطالب من الطلاب المسجلين في جـامعة الملك سعود، وليكن:

X = 1 أو 0 وفقا لما إذا كان يسكن أو لا يسكن في المدينة الجامعية.

Y = عدد إخوته.

Z = طوله بالسنتمتر.

فالمتغيرات X ، Y ، Z هي متغيرات عشوائية. ونلاحظ أن فضاء العينة لمثل هذه التجربة هو مجموعة الطلاب المسجلين في جامعة الملك سعود، كل طالب يمثل نقطة عينة (نتيجة ممكنة). وكل متغير من هذه المتغيرات الثلاثة يأخذ قيمة واحدة وواحدة فقط عند كل نقطة عينة: وهو من هذا الوجه يشكل دالة عددية معرفة على فضاء العينة. فمن أجل كل طالب يأخذ X قيمة واحدة فقط هي إما 1 أو 0، ويأخذ Y قيمة واحدة فقط هي عبارة عن عدد صحيح غير سالب. ويأخذ Z قيمة واحدة فقط هي عبارة عن عدد حقيقي موجب.

مثال (٣ ـ ٢)

لتكن التجربة هي قذف ثلاث قطع نقود، وليكن X عدد أوجه الـ H التي نحصل عليها. فالمتغير X هو متغير عشوائي قيمه الممكنة 0 أو 1 أو 2 أو 3. وهو يأخذ عند كل نقطة عينة من النقاط الثماني التي يتضمنها فضاء العينة لهذه التجربة قيمة واحدة فقط من هذه القيم الممكنة. والجدول (٣ ـ ١) يبين ذلك.

فضاء العينة S	(HHH)	(HHT)	(HTH)	(THH)	(HTT)	(THT)	(TTH)	(TTT)
X	3	2	2	2	1	1	1	0

ومن الواضح أن X يمثل دالة عددية معرفة على فضاء العينة S. $S \xrightarrow{X} \{0, 1, 2, 3\}$
وأن $X(HHH) = 3$ ، $X(HHT) = 2$ ، $X(HTH) = 2$ ، الخ . . .
ومما سبق يتضح لنا، بصورة عامة، التعريف التالي للمتغير العشوائي.

(٣ ـ ١ ـ ١) تعريف المتغير العشوائي

المتغير العشوائي هو دالة عددية معرفة على فضاء عينة.

وقد رأينا في الفصل السابق أن الحادثة هي مجموعة جزئية من فضاء عينة، فما هو حكم $X = 2$ ، مثلا، وهل تمثل حادثة؟ والجواب نعم لأن $X = 2$ تعني وقوع واحدة من النقاط (HHT) أو (HTH) أو (THH). وسنصطلح على أن $[X = 2]$ تمثل الصورة العكسية لـ $X = 2$ أو $X^{-1}(2)$ ونكتب:

$$[X = 2] = X^{-1}(2) = \{(HHT), (HTH), (THH)\}$$

وهذا يسمح لنا بالقول إن $X = 2$ حادثة عددية نعبر عنها بدلالة المتغير العشوائي. ذلك لأن لها ما يقابلها في فضاء العينة الأصلي S. ونلاحظ أكثر من ذلك أن الحوادث العددية $X = 0$ ، $X = 1$ ، $X = 2$ ، $X = 3$ هي حوادث متنافية وتشكل تجزئة لفضاء العينة الأصلي S ، ففي الواقع:

$X = 0$ تمثل الحادثة $X^{-1}(0) = \{(TTT)\}$ أي مجموعة نقاط العينة التي يأخذ فيها X القيمة 0،

$X = 1$ تمثل الحادثة $X^{-1}(1) = \{(HTT), (THT), (TTH)\}$ ، أي مجموعة نقاط العينة التي يأخذ فيها X القيمة 1 ،

$X = 2$ تمثل الحادثة $X^{-1}(2) = \{(THH), (HTH), (HHT)\}$ ، أي مجموعة نقاط العينة التي يأخذ فيها X القيمة 2 ،

$X = 3$ تمثل الحادثة $X^{-1}(3) = \{(HHH)\}$ ، أي مجموعة نقاط العينة التي يأخذ فيها X القيمة 3 .

وهي في مثالنا هنا، وفي غيره أيضا، متنافية بالضرورة، لأنه لو كان بين أي اثنين منها نقطة عينة مشتركة، لاقتضى ذلك أن يكون لـ X قيمتان مختلفتان في تلك النقطة، مما يتناقض مع حقيقة أن X دالة كما ينص التعريف. وسنقول إن المتغير X ولد فضاء عينة جديدا هو مجموعة قيمه الممكنة $\{0, 1, 2, 3\}$.

مثال (٣ ـ ٣)

نقذف قطعة نقود حتى يظهر وجه الـ H للمرة الأولى. وليكن X عدد القذفات التي نحتاجها. النتائج الممكنة للتجربة أو فضاء العينة هو:

$$H, TH, TTH, TTTH, \ldots$$

ومن الواضح أن X يمكن أن يكون 1 أو 2 أو 3 الخ. . . أي أن فضاء العينة الذي ولده X ، أو مجموعة قيمه الممكنة هي مجموعة الأعداد الطبيعية $\{1, 2, 3, 4, \ldots\}$.

(٣ ـ ٢) تصنيف المتغيرات العشوائية

ولنعد إلى المثال (٣ ـ ١)، ولنتساءل عن مجموعة القيم الممكنة لـ Z ، طول الطالب. بما أننا سنستخدم مسطرة مدرجة لقياس الطول فإن طول الطالب سيقابل

نقطة على هذه المسطرة هي، في الواقع، نقطة على محور موجه. والقيمة التي يأخذها Z يمكن أن تكون أي نقطة من فترة على محور موجه. وبالطبع يوجد في أي فترة من محور موجه، مهما كانت صغيرة، ما لا نهاية له ولا يمكن عده أو حصره من النقاط. وبالرغم من أن فضاء العينة الذي يولده المتغير X في المثال (٣ ـ ٣) لا نهائي أيضا. إلا أن هناك خلافا أساسيا بين طبيعتي الفضائين. فنقاط فترة من محور الأعداد الحقيقية هي مجموعة لا نهائية لا يمكن عدها، أي لا يمكن إقامة تقابل بين هذه النقاط وبين الأعداد الصحيحة الموجبة ... ,3 ,2 ,1. ولو أخذنا الفترة [160, 200]، مثلا، واعتبرنا 160 مقابلا للعدد الصحيح 1، ثم سألنا أنفسنا ما هو العدد الذي يليه أي العدد الذي سيقابل 2 لاستحالت الإجابة. ومهما كان العدد الذي نرشحه قريبا من 160 فسيبقى بين مثل هذا العدد والـ 160 ما لا يحصى ولا يعد من الأعداد. أما قابلية العد في فضاء العينة المتولد عن المتغير X في المثال (٣ ـ ٣) فهي أمر واضح لا يحتاج إلى تعليق. وهكذا نجد أن قابلية العد تميز بين صنفين من فضاءات العينة سنعرفهما فيما يلي:

(٣ ـ ٢ ـ ١) الفضاء المنفصل

نقول عن فضاء عينة إنه فضاء منفصل إذا كان يحوي عددا منتهيا من النقاط أو لا نهاية قابلة للعد من النقاط.

(٣ ـ ٢ ـ ٢) الفضاء المتصل

نقول عن فضاء عينة إنه فضاء متصل (أو مستمر) إذا كان يحوي لا نهاية غير قابلة للعد من النقاط.

ووفقا لهذا التصنيف نصنف المتيغرات العشوائية إلى متغيرات منفصلة ومتغيرات متصلة (أو مستمرة).

(٣ ـ ٢ ـ ٣) المتغير العشوائي المنفصل

نقول إن المتغير العشوائي منفصل إذا كانت مجموعة قيمه الممكنة مجموعة منتهية أو لا نهائية قابلة للعد. أي إذا كان فضاء العينة الذي يولده هذا المتغير فضاء منفصلا.

(٣ ـ ٢ ـ ٤) المتغير العشوائي المتصل (المستمر)

نقول إن المتغير العشوائي متصل (أو مستمر) إذا كانت مجموعة قيمه الممكنة لا نهائية وغير قابلة للعد. أي إذا كان فضاء العينة الذي يولده هذا المتغير متصلا (أو مستمرا).

(٣ ـ ٣) المتغيرات العشوائية المنفصلة وتوزيعاتها الاحتمالية

رأينا أنه يمكن التعرف على نوع المتغير العشوائي من خلال اتصاف مجموعة قيمه الممكنة أو عدم اتصافها بقابلية العد. وإذا أمكن للمتغير أن يفترض أو يتخذ عددا منتهيا أو لا نهائيا قابلا للعد من القيم فهو متغير منفصل. وفي معظم المسائل التي نواجهها في الحياة العملية تمثل المتغيرات المنفصلة قياسات على شكل تعداد مثل عدد حوادث المرور في مدينة الرياض خلال أسبوع، وعدد الإشارات الحمر التي تواجهها في طريقك إلى عملك، وعدد القطع المعيبة صناعيا في الإنتاج اليومي لمصنع، وعدد حالات الطلاق خلال سنة في مدينة معينة، وعدد البكتريا في سنتمتر مكعب من الماء، وعدد الطائرات التي تصل في اليوم في رحلات دولية إلى مطار الملك خالد الدولي إلخ. وإذا كان عدد الإشارات التي تجتازها في طريقك إلى عملك هو عشر إشارات فإن عدد الإشارات الحمر التي يمكن أن تواجهها يتراوح بين 0 و 10 وعدد البكتريا X في سنتمتر مكعب من الماء يمكن أن يكون كبيرا جدا إلا أنه محدود على أي حال، أي أن $X = 0, 1, \ldots, n$ حيث n عدد كبير جدا.

ودالة التوزيع لمتغير عشوائي منفصل هي صيغة أو جدول يعرض القيم الممكنة والاحتمال الموافق لكل قيمة.

مثال (٣ ـ ٤)

في المثال (٣ ـ ٢) أوجد التوزيع الإحتمالي لـ X .

الحـــل

بالعودة إلى فضاء العينة الأصلـي للتجربة وهو الفضاء المذكور في الجدول (٤ ـ ١) ومن اتزان أو تناظر قطع النقود، يمكننا إقامة نموذج احتمالي على هذا

الفضاء بتوزيع حصص متساوية على النقاط الثماني التي يتضمنها فضاء العينة . وبذلك يكون الاحتمال الموافق لكل نقطة هو 1/8 . وبعد أن نبني نموذجا احتماليا على فضاء العينة الأصلي، يمكننا الإجابة عن احتمال أي حادثة في هذا الفضاء (أي مجموعة جزئية من هذا الفضاء). ولكننا هنا في صدد الإجابة عن حادثة عددية معبر عنها بدلالة المتغير العشوائي X . مثلا، ما احتمال أن يأخذ X القيمة واحد. ونكتب ذلك رمزيا $P(X = 1)$ ، لنصطلح على ما تمليه البداهة هنا. وهو أن هذا الاحتمال هو احتمال الحادثة في فضاء العينة الأصلي التي تمثلها عبارة $X = 1$ ، أي احتمال الحادثة $X^{-1}(1) = \{(HTT), (THT), (TTH)\}$ وهو كما نعلم مجموع احتمالات النقاط الثلاث التي تتضمنها هذه الحادثة. أي $3/8 = 1/8 + 1/8 + 1/8$. وباختصار أكثر نقول إن احتمال $X = 1$ هو مجموع احتمالات نقاط العينة التي افترض فيها X القيمة 1 . وبتطبيق هذه القاعدة على بقية القيم الممكنة نجد الجدول (٣ ـ ٢) حيث $f(x) = P(X = x)$.

جدول (٣ ـ ٢)

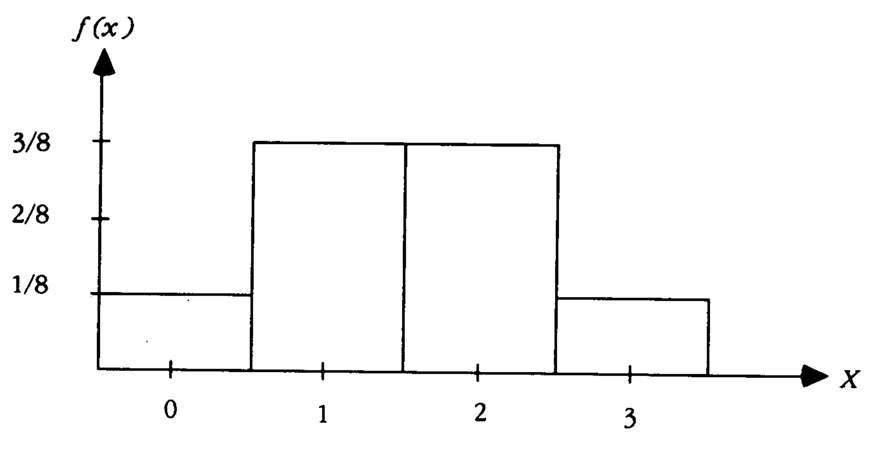
دالة التوزيع الاحتمالي لعدد أوجه الـ H عند قذف ثلاث قطع نقود.

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

ونلاحظ أولا أن مجموع الاحتمالات في هذا الجدول تساوي الواحد تماما. وهذه النتيجة متوقعة طالما أن الحوادث العددية التي تمثلها القيم المختلفة لـ X هي حوادث متنافية وتشكل تجزئة لفضاء العينة الأصلي S ، كما أوضحنا في المثال (٣ ـ ٢). وفي هذا المثال يمكننا تلخيص الجدول (٣ ـ ٢) بصيغة (علاقة) هي :

$$f(x) = \frac{\binom{3}{x}}{8} \ , \ x = 0, 1, 2, 3$$

ويمكن تمثيل هذا التوزيع بيانيا لنحصل على ما يسمى بالمدرج الاحتمالي. فلنتخذ القيم الممكنة مراكز لفترات تمتد بمقدار الواحد (نصف على يمين القيمة ونصف على يسارها) ولنرسم فوق كل فترة مستطيلا ارتفاعه يساوي الاحتمال الموافق فنحصل على مدرج الاحتمال كما في الشكل (٣ ـ ١).



شكل (٣ ـ ١) المدرج الاحتمالي للتوزيع في المثال (٣ ـ ٤).

وبصورة عامة، عندما نبني نموذجا احتماليا على فضاء عينة S يمكننا استنتاج دالة التوزيع الاحتمالي لأي متغير عشوائي، X مثلا، معرف على S. وذلك وفقا للقاعدة التالية:

$f(x) = P(X=x) =$ مجموع احتمالات نقاط العينة التي أخذ فيها X القيمة x

والتوزيع الاحتمالي لمتغير عشوائي منفصل ليس إلا نموذجا احتماليا نقيمه على فضاء العينة الذي ولده هذا المتغير العشوائي. وإذا تذكرنا الشروط التي يجب أن يحققها نموذج احتمالي كما وردت في الفقرة (٢ ـ ٩) يمكن أن نستنتج هنا القاعدة التالية:

يجب أن تحقق دالة التوزيع $f(x)$ لمتغير عشوائي منفصل X الشرطين التاليين:

١ ـ $f(x) \geq 0$ مهما تكن x.

٢ ـ $\sum_x f(x) = 1$ حيث $\sum_x$ تعني المجموع فوق جميع القيم الممكنة x.

مثال (٣ ـ ٥)

أوجد التوزيع الاحتمالي للمتغير X في المثال (٣ ـ ٣).

ومن هذا النموذج المعطى في الجدول (٣ ـ ٣) نستنتج بسهولة، وبتطبيق القاعدة العامة المعطاة أعلاه، دالة التوزيع الاحتمالي للمتغير X كما في الجدول (٣ ـ ٤). ويمكن التعبير عن هذه الدالة بالصيغة التالية:

الحل

جدول (٣ ـ ٤)

دالة التوزيع الاحتمالي لـ X

$f(x)$	X
(1/2)	1
$(1/2)^2$	2
$(1/2)^3$	3
$(1/2)^4$	4
:	:
:	:
:	:

جدول (٣ ـ ٣)

النموذج الاحتمالي على فضاء العينة الأصلي

الاحتمال الموافق	نقطة العينة
1/2	H
1/4	TH
1/8	TTH
1/16	$TTTH$
:	:
:	:
:	:

$$f(x) = \left(\frac{1}{2}\right)^x . \quad x = 1, 2, 3, \ldots$$

وللتحقق من أن الدالة $f(x)$ تحقق شرطي دالة التوزيع المذكورين أعلاه ، نلاحظ أولا أن جميع قيم الدالة غير سالبة وأن

$$\sum_x f(x) = \sum_{x=1}^{\infty} \left(\frac{1}{2}\right)^x = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \ldots$$

$$= \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1 .$$

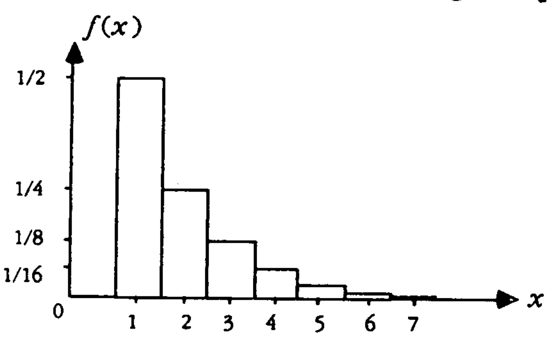
مثال (٣ ـ ٦)

ليكن X عدد النقاط على الوجه الظاهر عنـد رمي حجر نرد متماثل (متناظر) . ما هي دالة التوزيع الاحتمالي لـ X ؟

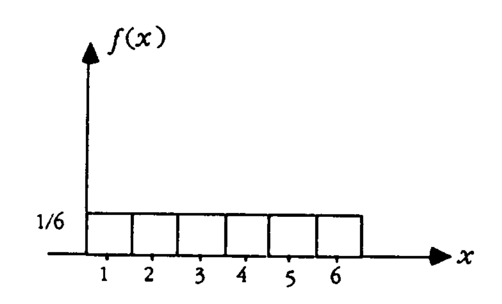
ويمكن التعبير عن الدالة هنا بالصيغة التالية :

$$f(x) = \frac{1}{6} \; , x = 1, 2, 3, 4, 5, 6.$$

ومن الواضح أن $f(x)$ تحقق شرطي دالة التوزيع. ويسمى مثل هـذا التوزيع بالتوزيع المنتظم لأن الواحد مـوزع بالتساوي (بانتظام) على كافـة القيم الممكنة لـ X . والمدرج الاحتمالي مرسوم في الشكل (٣ ـ ٣).



شكل (٣ ـ ٢) المدرج الاحتمالي للتوزيع في المثال (٣ ـ ٥)



شكل (٣ ـ ٣) المدرج الاحتمالي للتوزيع في المثال (٣ ـ ٦)

(٣ ـ ٤) التفسير العملي للتوزيع الاحتمالي لمتغير عشوائي منفصل

سنتصور مع كل متغير عشوائي، X مثلا، مجتمعا من القيـاسات، هو المجتمع الناشيء عن تكرار قياس X عددا هائلا (لا نهائيا) من المرات. والتوزيع الاحتمالي لـ X يقدم وصفا لبنية هذا المجتمع. ففي المثال (٣ ـ ٤) يقول لك التوزيع الاحتمالي:

لو أنك كررت تجربة قذف ثلاث قطع نقود متماثلة (متناظرة) بلا حدود، وفي كل تكرار سجلت قيمة X (عدد أوجه الـ H التي حصلت عليها) فسيقع ذلك المجتمع من

القياسات الذي تحصل عليه في أربع فئات هي، فئة الصفر، وفئة الـ 1 ، وفئة الـ 2 ، وفئة الـ 3 . ولو قمت بتصنيف وكتابة جدول التكرار النسبي لهذا المجتمع من القياسات، فستجد أن التكرار النسبي لكل من فئتي الصفر والواحد هو 12.5% ، وأن التكرار النسبي لكل من فئتي الـ 2 والـ 3 هو 37.5% . وبعبارة أخرى، لو أنك رسمت مدرج التكرار النسبي لهذا المجتمع من القياسات فستحصل على الصورة نفسها التي يقدمها لك المدرج الاحتمالي لتوزيع X . وفي المثال (٣ ـ ٥) يقول لك التوزيع الإحتمالي أنك لو كررت التجربة بلا حدود وسجلت في كل مرة عدد القذفات التي احتجت إليها حتى ظهور وجه الـ H للمرة الأولى، فستجد في 50% من هذه التكرارات أنك احتجت لقذفة واحدة، وفي 25% من التكرارات لقذفتين، وفي 12.5% من التكرارات لثلاث قذفات، وفي 6.25% لأربع قذفات، وهكذا . . . . وفي المثال (٣ ـ ٦) يقول لك التوزيع الاحتمالي أنك لو كررت تجربة رمي حجر نرد عددا لا نهائيا من المرات، فستظهر الأوجه الستة بتكرارات نسبية متساوية، وكل منها يساوي $16\frac{2}{3}\%$.

وبالطبع فإن تكرار أي تجربة عددا لا نهائيا من المرات هو مجرد افتراض نظري، أي أن المجتمع من القياسات الموافق لمتغير عشوائي ليس إلا مجتمعا تصوريا قائما في الذهن فقط. وفي الحقيقة ليس التوزيع الاحتمالي إلا تجريدا ذهنيا لحالة فيزيائية واقعية، أي أنه يشكل نموذجا رياضيا، ويقدم وصفا لمجتمع نظري بلغة الواقع، لغة الإحصاء الوصفي التي قدمناها في الفصل الأول من هذا الكتاب. ويكون المدرج الاحتمالي، بهذا المعنى، هو مدرج التكرار النظري لمجتمع القياسات.

ويبرز هنا سؤال جوهري. إذا كيف نتحقق من أن هذه التجريدات الذهنية تقدم محاكاة ناجحة لعالم الواقع؟

وللإجابة عن هذا السؤال يمكننا، كما هو الحال في العلوم التجريبية، اللجوء إلى التجربة والمشاهدة. وإذا كان توليد مجتمع لا نهائي غير ممكن عمليا إلا أنه يمكن تكرار التجربة عددا كبيرا من المرات ثم تصنيف القيم التي نحصل عليها للمتغير العشوائي ثم مقارنة صورة مدرج التكرار النسبي بصورة المدرج الاحتمالي. ولو قمنا بذلك لرأينا

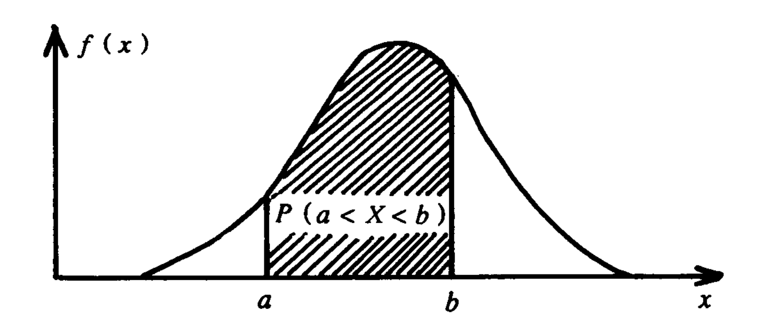
شبها يثير الإعجاب، حتى في حالـة أعداد معتدلة لتكرارات التجربـة. وهو شبه يزداد حـدة ووضوحـا مع زيادة عـدد التكرارات. ويمكن للقـارىء أن يقوم بتكـرار أي من التجارب المذكورة في الأمثلـة (٣ ـ ٤)، (٣ ـ ٥)، (٣ ـ ٦)، مـائة مرة، مثلا، ليحصل على مائة قياس للمتغير العشوائي X الذي نقيسه، ويرسم لهذه القياسات المائة مدرج تكرار نسبي يقارنـه بالمدرج الاحتمالي لـ X. وإذا كانت درجة التشـابه بينهما لا ترضيه، يمكنه زيادة عدد التكرارات مائة أخرى ورسم مدرج تكرار نسبي للمئتين من القياسات المتوفرة، وسيلاحظ أن الصورة الجديـدة لمدرج التكرار النسبي قد اعتدلت في اتجاه المزيد من الشبه بين المشاهدة التجريبية والمقال النظري.

وعندما نقف بعد n من التكرارات ننظر إلى العدد المحـدود من القياسات، n، على أنه عينة عشـوائية من مجتمع القياسـات الذي كنا سنحصل عليه لـو استمر تكرار التجربة بلا حدود، وهذه المقولـة هي مقولة إصطلاحية في علم الإحصاء ولها فوائد جمة في التطبيقات العملية.

(٣ ـ ٥) المتغيرات العشوائية المتصلة

تشكل الكميات التي نستخدم للحصول على مقاديرها أجهزة قياس، أو أدوات قياس، متغيرات عشوائيـة متصلة. فالوزن والقـوة والطول ومعدل هطول المطـر ودرجة حـرارة جسم ودرجـة الامتحان لطـالب كلهـا أمثلـة على متغيرات عشوائيـة مستمـرة. وقياسات مثـل هذه المتغيرات هي نقاط على خط اتخذنا عليـه تدريجا أو سلما للقياس، أي أنها نقاط على المحور الموجه (محور الأعداد الحقيقية)، أو على فترات من هذا المحور. ولا يمكننا، في حالة متغير عشـوائي مستمر، تخصيص أي احتمال مهما كان صغيرا لأي قيمة من قيم المتغير نظرا للكثـرة الكاثرة من القيم المختلفة، إذ توجد لا نهايـة غير قابلة للعد مـن النقاط في أي فترة مهما صغـرت، مما سيـؤدي إلى الخروج على مسلمة الاحتمال الثانيـة. ولا بد من التفكير في طـريقة لبناء النمـوذج الاحتمالي مختلفة تمامـا عما رأيناه في حالة متغير عشوائي منفصل.

لنعد بذاكرتنا الآن إلى مناقشة المضلع التكراري في الفقرة (١ ـ ٢ ـ ٣) ، حيث رأينا إمكانية تفسير المساحة تحت مدرج التكرالر النسبي كاحتمال. وإلى منحنى التكرار في الفقرة (١ ـ ٤) ، حيث تمثل كل نقطة على محور السينات قياسا ويمثل الإحداثي الصادي لتلك النقطة تواتر أو تكرار ظهور هذا القياس في المجتمع من القياسات الذي يصفه منحنى التكرار. إذ تقدم لنا هذه الأفكار نقطة البداية في محاولة بناء نموذج احتمالي لمتغير عشوائي مستمر. لنبدأ بالقول إنه إذا كان تكرار ظهور القياس a ، مثلا، أكبر من تكرار ظهور القياس b ، فإن الكثافة الاحتمالية في a ينبغي لها أن تكون أكبر من الكثافة الاحتمالية في b . ولنعتبر منحنى التكرار منحنى كثافة يبين لنا كيف تتغير الكثافة الاحتمالية من نقطة إلى أخرى. ولنسمي الدالة المستمرة $f(x)$ ، التي بيانها هو منحنى التكرار، دالة كثافة احتمالية. وعندئذ ستمثل المساحات تحت هذا المنحنى احتمالات. واحتمال أن يقع قياس المتغير x ضمن فترة (a, b) أي $P(a < X < b)$ هو المساحة تحت منحنى الكثافة وفوق الفترة (a, b) . (انظر الشكل ٣ ـ ٤).



شكل (٣ ـ ٤) دالة الكثافة الاحتمالية $f(x)$ أو منحنى التكرار

أي أن $P(a < X < b)$ يساوي قيمة الدالة الأصلية لـ $f(x)$ محسوبة عند b مطروحا منها قيمة الدالة الأصلية عند a .

وترتب علينا مثل هذه الطريقة شرطين، لا بد لأي دالة كثافة أن تحققهما، كي لا نخرج على مسلمات الاحتمال. فما دام الاحتمال غير سالب، لا يجوز أن يكون أي

جزء من منحنى الكثافة تحت المحور السيني. وبما أن احتمال الحادثة الأكيدة، أي $P(-\infty < X < +\infty)$، يجب أن يكون مساويا للواحد تماما، فإن المساحة تحت منحنى الكثافة بكامله يجب أن تساوي الواحد تماما. وهكذا نكتب القاعدة التالية:

(٣ـ٥ـ١) قاعدة

كي تصلح دالة متصلة $f(x)$ كدالة كثافة احتمالية يجب أن تحقق الشرطين التاليين:

١ ـ $f(x) \geq 0$ مهما يكن x،

٢ ـ المساحة تحت بيان $f(x)$ (أي تحت منحنى الكثافة) تساوي الواحد تماما.

ووفقا لهذا التصور لو سألنا الآن عن احتمال أن يفترض متغير عشوائي متصل X قيمة محددة x، مثلا لكان الجواب:

$P(X = x) =$ المساحة تحت منحنى الكثافة فوق النقطة x $= 0$

والاحتمال صفر يعني الاستحالة. وهنا نجد أنفسنا في مأزق. لنفرض، للتوضيح، أن X يمثل طول إنسان ذكر بالغ بالسنتمتر. فاستحالة أن يكون هذا المتغير مساويا لقيمة محددة، أي قيمة، تعني نفي وجود الجنس البشري، وهي نتيجة في غاية السخف، مما يثير الريبة في صلاحية النموذج الرياضي الموضوع لمتغير عشوائي متصل. ولكن لو تأملنا قليلا في هذه النتيجة لوجدنا أن التفسير المنطقي الوحيد لها هو أنه يستحيل على الإنسان أن يبتكر جهازا للقياس لا يخطىء، أو أن يقيس بدون خطأ. ولا ريب أن لطول إنسان ذكر بالغ قيمة محددة تماما، والمستحيل ليس وجود هذا القيمة وإنما القدرة على معرفتها، أي أن ما يستحيل هنا هو الادعاء بأننا نستطيع قياس الأطوال بدون خطأ. ويجدر أن نقف قليلا أمام هذا المثال لنجد كيف يضطر المكابرون لتسجيل عجز الإنسان أمام بارئه في شكل معادلة رياضية، وفي ذلك آية لذوي البصيرة.

ويبقى سؤال وجيه آخر، إذ كيف نختار النموذج الموافق لحالة معينة؟ وما يمكن قوله هنا هو أن نستفيد من كل المعلومات المتوافرة لنا ثم نختار النموذج $f(x)$ وفق أفضل ما لدينا من قدرة على الحكم الصحيح. وتتفرع المسألة هنا إلى مسألتين، فمثلا،

قد نعرف أن $f(x)$ على شكل جرس، ولكن من بين مثل هذه الأسرة من النماذج، ما هو على وجه التحديد، ذلك النموذج (ذلك المنحنى على شكل جرس) الذي يوافق الحالة المدروسة؟ وقد لا نعرف، على الوجه الآخر، حتى الشكل الأولي لـ $f(x)$، ونتساءل، مثلا، عما إذا كان ينبغي افتراضه على شكل جرس أم لا ؟ ويتطرق الاستقراء الإحصائي إلى كل من المسألتين. ويقدم لنا الإحصاء الرياضي طرقا لمعالجة مثل هذه المسائل سواء أكان المتغير X مستمرا أم منقطعا. وبعد أن يقع اختيارنا على النموذج المناسب يمكننا في حالة متغير متصل حساب أي مساحة تحت منحنى الكثافة باستخدام الحساب التكاملي، وفي العديد من النماذج المعروفة والمستخدمة على نطاق واسع في طرق الإحصاء تتوافر جداول جاهزة تزودنا بمثل هذه المساحات.

ولكن هل يمكن الحصول على نتائج مفيدة باستخدام نماذج لم نتأكد تماما من صحتها، أي من تمثيلها بصورة دقيقة للمجتمع المدروس؟ لننظر هنا إلى المهندس والكيميائي والفيزيائي وغيرهم، فنجد أن مختلف العلاقات العددية المستخدمة في مختلف فروع العلوم هي نماذج رياضية تقدم لنا تقريبات جيدة لواقع الحياة العملي. ويستخدم المهندس معادلاته لتحديد حجم وموضع دعامات جسر أو حجم وموضع جناح طائرة وما يهمه، في المقام الأول، هو أن تقدم الجسور وأجنحة الطائرات الخدمات التي صممت من أجلها. وبالمثل فإن ما تقدمه الطرق الإحصائية من خدمات، هو المسطرة التي نقيس بها فائدة هذه الطرق، والقاعدة التي نحكم من خلالها على صحة ما تزودنا به من تنبؤات وقرارات تتعلق بالمجتمع المدروس. والجواب على سؤالنا نجده بوضوح في تطبيقات الإحصاء التي عم استخدامها وثبتت فوائدها، وقد اتسعت مساحة هذه التطبيقات لتشمل، على وجه التقريب، مختلف ميادين المعرفة ولتصبح بحق أداة رئيسة من أدوات الإنسان المعاصر في سعيه الدائم للكشف عن المجهول تحت شروط خاضعة للمصادفة.

(٣ ـ ٦) دالة التوزيع الاحتمالي المتجمع

رأينا في الفقرة (١ ـ ٣) أن التكرار المتجمع الصاعد يجيب عن السؤال التالي: ما هو التكرار النسبي لظهور قياس يقل عن قيمة محددة؟ وستجيب دالة التوزيع المتجمع

عن سؤال مشابه: ما احتمال أن يأخذ متغير عشوائي X قيمة أقل من أو تساوي قيمة محددة؟ وإذا رمزنا لهذه الدالة بـF فإن قيمة F في نقطة x هي ببساطة احتمال أن يأخذ المتغير X قيمة أقل من أو تساوي x، أي $P(X \leq x)$.

(٣ـ٦ ـ ١) حالة متغير عشوائي منفصل

دالة التوزيع المتجمع لمتغير عشوائي منفصل X، دالة احتماله $f(x)$ هي بالتعريف:

$$F(t) = \sum_{x \leq t} f(x)$$

حيث $\sum_{x \leq t}$ تعني المجموع فوق جميع القيم الممكنة التي تقل عن t أو تساويها.

مثال (٣ـ٧)

في المثال (٣ ـ ٤) ما احتمال الحصول على وجه الـ H مرتين على الأكثر؟

الحل

المطلوب هو حساب

$$\begin{aligned} P(X \leq 2) = F(2) &= \sum_{x=0}^{2} f(x) \\ &= f(0) + f(1) + f(2) \\ &= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8} \end{aligned}$$

تمرين

اعط تفسيرا عمليا لهذه النتيجة.

مثال (٣ـ٨)

في المثال (٣ ـ ٥) ما احتمال ألا يحتاج ظهور وجه الـ H للمرة الأولى إلى أكثر من ثلاث قذفات؟

الحل

المطلوب هو حساب :

$$P(X \le 3) = F(3) = \sum_{x=1}^{3} f(x)$$
$$= f(1) + f(2) + f(3)$$
$$= \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 = \frac{7}{8}$$

تمرين

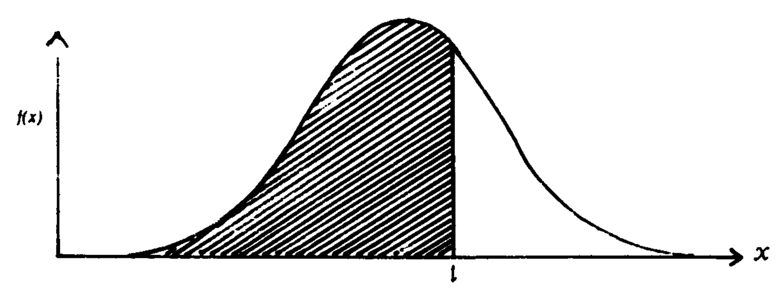
اعط تفسيرا عمليا لهذه النتيجة .

(٣ ـ ٦ ـ ٢) حالة متغير عشوائي متصل

دالة التوزيع المتجمع لمتغير عشوائي متصل X دالة كثافته الاحتمالية $f(x)$ ، هي بالتعريف :

المساحة تحت منحني الكثافة الواقعة إلى اليسار من النقطة t = $P(X \le t) = F(t)$

انظر الشكل (٣ ـ ٥) .



شكل (٣ ـ ٥) : دالة التوزيع المتجمع لمتغير عشوائي متصل .

قلنا إن لكل متغير عشوائي مجتمعا من القياسات هو المجتمع الناشيء عن تكرار قياس المتغير العشوائي مرة بعد أخرى إلى ما شاء الله . وأن التوزيع الاحتمالي يقدم وصفا للبنية الداخلية لهذا المجتمع ويمثل التوزيع التكراري النظري له . وكما أن لكل مجموعة من القياسات مقاييس للنزعة المركزية ومقاييس للتشتت فكذلك الأمر بالنسبة إلى مجتمع القياسات . وسنتحدث في الفقرتين القادمتين عن متوسط مجتمع القياسات وعن

تباينه، على الترتيب. وسنصطلح على استخدام عبارة «متوسط المجتمع» أو عبارة «متوسط التوزيع» لتعني الشيء نفسه. وكذلك سنقول في الوقت نفسه «تباين المجتمع» أو «تباين التوزيع»، و«الانحراف المعياري للمجتمع» أو «الانحراف المعياري للتوزيع». وسنرمز، كما جرت العادة في أدبيات الإحصاء، لمتوسط مجتمع بالحرف اليوناني μ (ننطقه «ميو»)، وللانحراف المعياري لمجتمع بالحرف اليوناني σ (ننطقه «سيجما»).

(٣ ـ ٧) التوقع الرياضي

(٣ ـ ٧ ـ ١) التوقع الرياضي لمتغير X

ليكن X متغيرا عشوائيا منفصلا دالة احتماله $f(x)$. ولنرمز لتوقع X بـ $E(X)$، فعندئذ:

$$E(X) = \sum_x x \,.\, f(x)$$

حيث $\sum_x$ يعني المجموع فوق كل القيم الممكنة للمتغير X.

هذا التعريف يقدم قاعدة لحساب توقع متغير عشوائي منفصل؛ إذ نضرب كل قيمة من القيم الممكنة للمتغير بالاحتمال المقابل لها ونجمع الجداءات الناتجة فنحصل على ما يسمى «بالقيمة المتوقعة رياضيا للمتغير»، أو اختصارا «القيمة المتوقعة للمتغير».

وكنا رأينا في الفقرة (١ ـ ٦ ـ ١) أن متوسط بيان مبوب يتضمن n قياسا هو:

$$\bar{y} = \frac{\sum_{i=1}^{m} y_i f_i}{n}$$

$$= y_1 \frac{f_1}{n} + y_2 \frac{f_2}{n} + \ldots + y_m \frac{f_m}{n}$$

حيث $y_m, \ldots, y_1$ هي القيم المختلفة التي يتضمنها البيان و f_1/n هو التكرار النسبي لـ y_1 و f_2/n هو التكرار النسبي لـ y_2 وهكذا. أي أنه لحساب المتوسط نضرب كل قيمة

بالتكـرار النسبي لظهور هذه القيمـة في البيان الإحصـائي ثم نجمع الجداءات الناتجة. وبالعـودة إلى التفسير العملي لدالـة التوزيع الاحتمالي يتضـح لنا أن $E(X)$ هو متوسط مجتمع القياسات. فنحن في عبارة $E(X)$ إنما نضرب كل قيمة من القيـم المختلفة التي يتضمنها مجتمع القياسات بـ $f(x)$ الذي يمثل التكرار النسبي لظهـور هذه القيمة في مجتمع القياسـات. ويصبح المعنى التطبيقي لـ $E(X)$ أو التفسير العملي له واضحـا. فالقيمة المتوقعة $E(X)$ للمتغير X هي متوسط القيم التي يمكن أن يفترضها هذا المتغير على المدى الطويل، أي بعبارة أخرى متوسط مجتمع القياسات الموافق للمتغير X .

وتسمية $E(X)$ بالتوقع الرياضي ينبغـي ألا تثير أي التباس إذ نستخدم في حسابه نموذجـا رياضيـا مجردا هو التوزيع الاحتمالي، وهـو يعبر، في الحقيقة، عن خـاصة من خواص هذا النموذج الرياضي. إذ تمثل قيمة $E(X)$ الموضع أو النقطة على محور السينات (محو الأعـداد) التي يتمركـز حـولها التوزيع الاحتمالي لــ X ، ولـذلك سنسميها أيضـا متوسط التوزيع الاحتمالي.

مثال (٣ ـ ٩)

في المثال(٣ ـ ٤) احسب $E(X)$.

الحل

$$E(X) = \sum_{x=0}^{3} x \cdot f(x)$$

$$= 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{12}{8} = 1.5$$

ومما سبق يمكن القول إن:

(i) 1.5 هي القيمة المتوقعة رياضيا لعدد أوجه الـ H .

(ii) يتمركز التوزيـع الاحتمالي لعدد أوجه الـ H حول النقطة 1.5 . ولو نظرنـا إلى صورة المـدرج الاحتمالي في الشكل (٤ ـ ١) لوجدنـا أنه يتمركز حول النقطة 1.5 على محور الـx . فالقيمة 1.5 هي متوسط التوزيع الاحتمالي.

(*iii*) التفسير العملي للقيمة 1.5 هو أنها تمثل القيمة المتوسطة لعدد أوجه الـ H على المدى الطويل (أي متوسط مجتمع القياسات). بمعنى أننا لو كررنا قذف قطع النقود الثلاث عددا هائلا من المرات وسجلنا في كل مرة عدد أوجه الـ H التي حصلنا عليها ثم حسبنا متوسط هذه الأعداد لحصلنا على 1.5 .

بعد أن تعلمنا كيفية حساب التوقع الرياضي لمتغير عشوائي X ، سندرس الآن طريقة حساب القيمة المتوقعة لدالة في X ، $g(X)$ مثلا. لنعد إلى المثال (٣ ــ ٤)، ولنفرض أن $g(X) = X^2$. من الواضح أن X^2 يأخذ عند كل نقطة عينة في الجدول (٣ ـ ١) قيمة واحدة وواحدة فقط، أي أنه دالة معرفة على فضاء عينة وبالتالي فهو متغير عشوائي. ويبين هذا، بصورة عامة، أن كل دالة في متغير عشوائي هي بدورها متغير عشوائي. ولكن كيف نحسب التوقع الرياضي لـ X^2 ؟ بما أن التوقع الرياضي يمثل متوسط مجتمع القياسات، فلنبحث عن كيفية حساب متوسط مجتمع القياسات ومنها نستنتج قاعدة لحساب التوقع الرياضي. ولكن ما هو مجتمع القياسات الموافق لـ X^2 ؟ إنه بالضبط مجتمع القياسات لـ X بعد تربيع كل قيمة من قيمه. وإذا كان ظهور القيمة $X = 2$ يتكرر بنسبة 3/8 ، كما نعلم من دالة التوزيع الاحتمالي لـ x ، فإن التكرار النسبي لظهور القيمة 4 في مجتمع القياسات الموافق لـ X^2 سيكون، في مثالنا هنا، 3/8 أيضا، وكذلك الأمر بالنسبة لبقية القيم. ولحساب القيمة المتوسطة، على المدى الطويل، نضرب كل قيمة ممكنة لـ X^2 بالتكرار النسبي لظهور هذه القيمة في مجتمع القياسات ثم نجمع النتائج، أي:

$$\text{متوسط مجتمع القياسات لـ } X^2 = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8}$$
$$= \frac{24}{8} = 3$$

وهذه النتيجة تسمح لنا بكتابة:

$$E(X^2) = \sum_{x=0}^{3} x^2 . f(x)$$

حيث $f(x)$ هي دالة التوزيع الاحتمالي للمتغير X .

وتقترح علينا المناقشة السابقة، بوضوح، التعريف التالي لتوقع دالة $g(X)$ ، بصورة عامة.

(٣ـ٧ـ٢) التوقع الرياضي لدالة عددية في X

ليكن X متغيرا عشوائيا منفصلا، دالة توزيعه الاحتمالي $f(x)$ ، ولتكن $g(X)$ دالة عددية في x فعندئذ:

$$E[g(X)] = \sum_{x} g(x) . f(x)$$

مثال (٣ـ١٠)

احسب $E(X^2)$ في المثال (٣ـ٦).

$$E(X^2) = \sum_{x=1}^{6} x^2 . f(x)$$
$$= 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^3 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6}$$
$$= \frac{91}{6} = 15\frac{1}{6}$$

وبصورة مشابهة تماما نعرف توقع متغير عشوائي مستمر. كل ما في الأمر أن دالة الكثافة الاحتمالية تقوم مقام دالة التوزيع الاحتمالي للمتغير المنفصل وتصبح إشارة المجموع $\sum$ إشارة تكامل $\int$ وسوف لا نتطرق لذلك في هذا الكتاب.

ومن خواص إشارة المجموع $\sum$ كما وردت في الفقرة (١ ـ ٥) يمكننا، بسهولة، التحقق من الخواص التالية لإشارة التوقع E .

(٣ـ٧ـ٣) خواص التوقع الرياضي

١ ـ إذا كان c عددا ثابتا [$g(X)$ في الفقرة (٣ ـ ٧ ـ ٢) تساوي مقدارا ثابتا c] فإن:

$$E(c) = \sum_{x} c . f(x) = c \sum_{x} f(x) = c$$

لماذا؟

٢ ـ إذا كانت $g(X) = cX$ حيث c عدد ثابت فإن :

$$E(cX) = \sum_x cx\, f(x) = c \sum_x x\, f(x) = c\,.\, E(X)$$

٣ ـ إذا كان $g(X) = g_1(X) + g_2(X)$ فإن

$$E[g(X)] = \sum_x g(x) \cdot f(x)$$

$$= \sum_x [g_1(x) + g_2(x)]\; f(x)$$

$$= \sum_x g_1(x) f(x) + \sum_x g_2(x)\; f(x)$$ (استنادا إلى خواص $\sum$)

$$= E[g_1(X)] + E[g_2(X)]$$

ومنه نستنتج الخاصة :

$$E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$$

وبصورة خاصة، إذا كان X_1 و X_2 أي متغيرين عشوائيين فإن :

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

٤ ـ من الخاصتين السابقتين يمكننا أن نكتب، بصورة عامة،

$$E[c_1 X_1 + c_2 X_2 + \ldots + c_n X_n]$$
$$= c_1 E(X_1) + c_2 E(X_2) + \ldots + c_n E(X_n)$$

حيث X_1، X_2، . . . ، X_n متغيرات عشوائية و c_1، c_2، . . . ، c_n أعداد ثابتة .

مثال (٣ ـ ١١)

تقدم الإحصائية التـالية وصفـا لمجتمع الأسر التي تقطن مـدنـا كبيرة من حيث خاصية امتلاكها للسيارات :

20% من الأسر لا تمتلك أي سيارة و 50% من الأسر تمتلك سيارة واحدة و 15% من الأسر تمتلك سيارتين و 10% من الأسر تمتلك ثلاث سيارات و 5% من الأسر تمتلك

أربع سيارات . إذا رمزنا بـ X لعدد السيارات وبـ Y لعدد العجلات التي تمتلكها أسرة .

ا ـ ما متوسط عدد السيارات للأسرة الواحدة؟

ب ـ أحسب متوسط عدد العجلات للأسرة الواحدة .

الحل

ا ـ من الواضح أن الوصف المعطى لمجتمع الأسر والمتعلق بقياسات X في هذا المجتمع يقدم دالة التوزيع الاحتمالي لـ X :

X	0	1	2	3	4
$f(x)$	0.20	0.50	0.15	0.10	0.05

ومتوسط عدد السيارات للأسرة الواحدة هو $E(X)$. ومن التعريف لدينا :

$$E(X) = 0(0.2) + 1(0.50) + 2\,(0.15) + 3\,(0.10) + 4\,(0.05) = 1.3$$

ب ـ عدد العجلات عند أسرة هو عـدد السيارات التي تمتلكها مضروبا بـ 5 . وإذا رمزنا لهذا المتغير العشوائي بـ Y فإن $Y = 5X$. والمطلوب هو $E(Y)$. ومن خواص التوقع لدينا :

$$E(Y) = E(5X) = 5E(X) = 5(1.3) = 6.5$$

ومتوسط عدد العجلات للأسرة الواحدة هو 6.5 عجلة .

ذكـرنا في مطلـع هـذه الفقـرة أن لمجتمع القيـاسات الموافق لمتغير عشـوائي X تباينـا وسنسمي مثـل هـذا التباين «تباين X» أو «تباين توزيع X». ونتذكر في الفقرة (١ ـ ٨ ـ ٤) أن تباين n من القياسات هو:

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

ولأغراض تتعلق بالاستقراء الإحصائي نقسم على $(n - 1)$ بدلا من n عندما نحسب تباين عينـة من القياسـات). وعنـدما نكتـب $\frac{1}{n}\sum_{1}^{n}\ldots$ فهذا يعني أننـا نجمع n مقدارا ثم نقسم على n ، أي نحسب متوسطا . ويمكن التعبير عن التباين كلاميا كما يلي :

تباين n من القياسات هو متوسط مربعات انحرافات القياسات عن متوسطها .

وستبنى العبارة نفسها في تباين مجتمع من القياسات موافق لمتغير عشوائي X ، فمن المعروف أن متوسط هذا المجتمع هو $E(X)$ ، وأن مربع انحراف قياس عن المتوسط هو $[X - E(X)]^2$ ، والتباين ما هو إلا توقع هذا المقدار (أي قيمته المتوسطة). ومنه التعريف التالي :

(٣ـ٧ـ٤) تباين متغير عشوائي

تباين متغير عشوائي X ، ونرمز له بـ $V(X)$ (أو σ_X^2) هو

$$V(X) = E[X - E(X)]^2$$

ومن خواص التوقع نستنتج الآن شكلا مختزلا أصلح للحسابات .
وبغية الاختصار سنكتب μ بدلا من $E(X)$ فنجد :

$$V(X) = E(X - \mu)^2 = E(X^2 - 2\mu X + \mu^2)$$

$$= E(X^2) - 2\mu E(X) + E(\mu^2)$$ (الخاصة الرابعة من خواص التوقع) ،

$$= E(X^2) - 2\mu^2 + \mu^2$$ (الخاصة الأولى من خواص التوقع) ،

$$= E(X^2) - \mu^2$$

ومنه الصيغة المختزلة للتباين :

$$V(X) = E(X^2) - [E(X)]^2$$

أو

$$V(X) = E(X^2) - \mu^2$$

أي أنه لحساب تباين X ، نحسب توقع مربع X ونطرح من الناتج مربع توقع X .

مثال (٣ـ١٢)

في المثال (٣ـ٤)، احسب تباين X .

$$V(X) = E(X^2) - \mu^2$$

ولكن

$$E(X^2) = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} = \frac{24}{8} = 3$$

ونعلم من المثال (٣ ـ ٩) أن $\mu = E(X) = 1.5$ ، إذن :

$$V(X) = 3 - (1.5)^2 = 3 - 2.25 = 0.75$$

(٣ ـ ٧ ـ ٥) الانحراف المعياري لمتغير

الانحراف المعياري لمتغير عشوائي X وسنرمز له بـσ_x ، أو اختصارا σ عندما نأمن الالتباس، هو الجذر التربيعي الموجب للتباين.

$$\sigma_X = \sqrt{V(X)}$$

وفي المثال السابق، الانحراف المعياري لعدد أوجه الـ H الناتجة عن قذف ثلاث قطع متزنة من النقود هو

$$\sigma = \sqrt{0.75} = 0.865$$

خواص التباين

١ ـ تباين العدد الثابت هو الصفر.

$$V(c) = E(c^2) - [E(c)]^2 = c^2 - c^2 = 0$$

٢ ـ $V(cX) = c^2 V(X)$ حيث X أي متغير عشوائي و c عدد ثابت.

$$V(cX) = E(cX)^2 - [E(cX)]^2$$
$$= c^2 E(X^2) - [cE(X)]^2$$
$$= c^2 [E(X^2) - (E(X)^2]$$
$$= c^2 V(X)$$

ونستنتج من هذه الخاصة أن :

$$\sigma_{cX} = |c| \sigma_X$$

أي إذا ضربنا المتغير X بعدد ثابت c فإن الانحراف المعياري لـ X يضرب أيضا بالعدد $|c|$.

٣ ـ يمكن البرهان أنه إذا كان المتغيران العشوائيان X_1 و X_2 مستقلين فيما بينهما فإن

$$V(X_1 + X_2) = V(X_1) + V(X_2)$$

ويمكن تعميم هذه القاعدة إلى أكثر من متغيرين فنقول إنه إذا كانت المتغيرات $X_1, X_2, \ldots, X_n$ مستقلة فيما بينها فإن :

$$V(X_1 + X_2 + \ldots, X_n) = V(X_1) + V(X_2) + \ldots + V(X_n)$$

ملاحظة

ما ذكرناه في الفقرة (١ ـ ٨) عن متباينة تشيبيشيف كان استعارة مبسطة لنظرية رياضية تحمل هذا الاسم وتتعلق بالتوزيعات الاحتمالية.

متباينة تشيبيشيف: ليكن X متغيرا عشوائيا متوسطه μ وانحرافه المعياري σ وليكن k أي عدد موجب، فعندئذ:

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

أي

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

أي أن احتمال أن يأخذ X قيمة تختلف عن المتوسط μ بأقل من k انحرافا معياريا هو على الأقل $1 - \frac{1}{k^2}$. وبلغة هندسية نقول في حالة متغير عشوائي منفصل إن $1 - \frac{1}{k^2}$ على الأقل من المساحة تحت مدرج الاحتمال واقع بين $\mu - k\sigma$ و $\mu + k\sigma$. وفي حالة متغير عشوائي مستمر نقول إن ما لا يقل عن $1 - \frac{1}{k^2}$ من المساحة تحت منحنى الكثافة الاحتمالية واقع بين $\mu - k\sigma$ و $\mu + k\sigma$.

مثال (٣ ـ ١٣)

في المثال (٣ ـ ١١) احسب تباين X وانحرافه المعياري، وتباين Y وانحرافه المعياري.

$$V(X) = E(X^2) - [E(X)]^2$$

وقد حسبنا في المثال (٣ ـ ١١) توقع X فوجدناه 1.3، ولدينا:

$$E(X^2) = 0^2 \times 0.20 + 1^2 \times 0.50 + 2^2 \times 0.15 + 3^2 \times 0.1 + 4^2 \times 0.05 = 3$$

ومنه:

$$V(X) = 3 - (1.3)^2 = 1.31$$
$$\sigma_X = \sqrt{V(X)} = \sqrt{1.31} = 1.1446$$

ومن الخاصة الثانية من خواص التباين نجد :

$$V(Y) = V(5X) = 25V(X) = 25 \times 1.31 = 32.75$$
$$\sigma_Y = \sqrt{32.75} = 5.72$$

أو

$$\sigma_Y = 5\sigma_X = 5(1.1446) = 5.723$$

تمارين (٣ـ١)

١) في كل مما يلي حدد ما إذا كانت الدالة f تصلح دالة توزيع احتمالي لمتغير عشوائي مجموعة قيمه الممكنة هي $\{1, 2, 3, 4\}$:

أـ $f(1) = 0.26,\ f(2) = 0.26,\ f(3) = 0.26,\ f(4) = 0.26$

ب ـ $f(1) = 0.15,\ f(2) = 0.28,\ f(3) = 0.29,\ f(4) = 0.28$

جـ ـ $f(1) = \frac{1}{9},\ f(2) = \frac{2}{9},\ f(3) = \frac{1}{3},\ f(4) = \frac{1}{3}$

د ـ $f(1) = 0.33,\ f(2) = 0.37,\ f(3) = -0.03,\ f(4) = 0.33$

هـ ـ $f(1) = \frac{1}{4},\ f(2) = \frac{3}{8},\ f(3) = \frac{3}{16},\ f(4) = \frac{5}{32}$

٢) حدد ما إذا كانت الدوال التالية تصلح دوال توزيع احتمالي وعلل إجابتك :

أـ $f(x) = \frac{1}{5},\quad x = 1, 2, 3, 4, 5;$

ب ـ $f(x) = \frac{x+1}{4},\ x = 1, 2, 3, 4;$

جـ ـ $f(x) = \frac{x^2}{30},\quad x = 1, 2, 3, 4;$

د ـ $f(x) = \frac{x-2}{5},\ x = 1, 2, 3, 4.$

وارسم المدرج الاحتمالي لكل دالة توزيع تجدها .

٣) حزمة من البطـاريات تتضمن 6 بطاريات ، اثنتان منها فاسـدتان . اخترنا عشوائيا عينة من ثـلاث بطاريـات . إذا رمزنـا بـ X لعدد البطـاريات الفاسـدة في العينة . أكتب التوزيع الاحتمالي لـ X ، وارسم المدرج الاحتمالي .

٤) يتضمن صنـدوق أربع قطع صـالحة وقطعة فـاسـدة . فحصنا هـذه القطع واحـدة فأخـرى . وليكن X رقم الاختبـار الـذي عثرنا فيـه على القطعة الفـاسدة . اكتب توزيع X .

٥) في كل من التمرينين (٣) و (٤) . أحسب متوسط التوزيع وتباينه .

٦) قذفنا حجري نرد؛ وليكن X عد النقاط الظاهرة على الحجر الأول ، و Y عدد النقاط الظاهرة على الحجر الثاني .

ا ـ اكتب التوزيع الاحتمالي لـ $T = X + Y$ واحسب $E(T)$ ، $V(T)$.

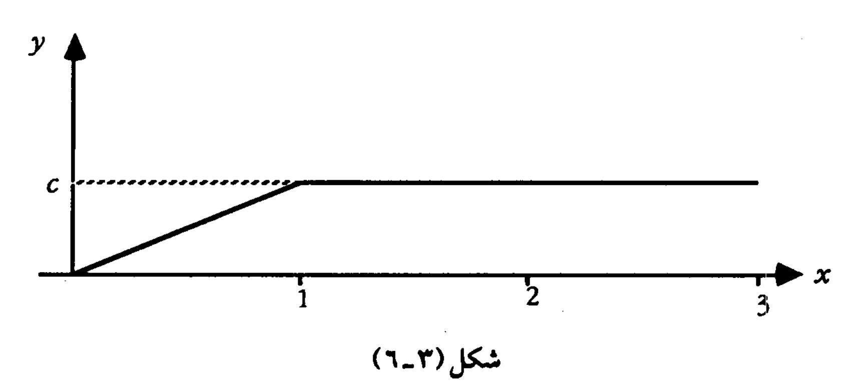
ب ـ اكتب التوزيع الاحتمالي لـ $W = XY$ ، واحسب $E(W)$ ، $V(W)$.

جـ ـ مستخـدما التوزيـع الإحتمـالي لكـل من X و Y ، احسب $E(X)$ ، $E(Y)$ ، $V(X)$ و $V(Y)$.

د ـ قارن بين $E(T)$ و $E(X) + E(Y)$ ؛ وبين $V(T)$ و $V(X) + V(Y)$.

٧) في الشكل (٣ ـ ٦) المجـاور، حدد قيمة c بحيث تصلح الدالـة المرسومة في الشكل دالة كثافة احتمالية ، واحسب :

$P(X < 1.5)$ ، $P(X = 2)$ ، $P(X \le 1.5)$ ، $P(X > 2.5)$ ، $P(0.5 < X < 2.5)$.



شكل (٣ـ٦)

٨) قذفنا ثلاث قطع نقود ، وليكن X عدد أوجه الـ H التي حصلنا عليها .
ا ـ اكتب التوزيع الاحتمالي لـ X ، وارسم المدرج الاحتمالي .
ب ـ احسب متوسط التوزيع $E(X)$ ، وتباين التوزيع $V(X)$.
جـ ـ نفذ هذه التجربة عمليا مائة مرة ، وسجل في كل مرة قيمة X ، ثم ارسم مدرج تكرار للقيم المائة لـ X . هل تجد أنه مشابه للمدرج الاحتمالي؟ استنتج من ذلك تفسيرا عمليا للمدرج الاحتمالي .
د ـ احسب $\bar{X}$ و S^2 متوسط وتباين القيم المائة لـ X التي حصلت عليها في جـ .
هل يشكل $\bar{X}$ تقديرا جيدا لـ $E(X)$ ، و S^2 تقديرا جيدا لـ $V(X)$ ؟

٩) مجتمع من خمسة أرقام هي $\{1, 2, 3, 4, 5\}$. سحبنا عشوائيا عينة من رقمين ، وليكن $\bar{X}$ متوسط هذه العينة . أكتب التوزيع الاحتمالي لـ $\bar{X}$.

١٠) تاجر للمعدات الثقيلة يتصل في اليوم بزبون واحد أو زبونين ، وذلك باحتمال يساوي 1/3 ، 2/3 ، على الترتيب . وسينتج كل إتصال إما لا شيء ، أو صفقة بيع قيمتها خمسون ألف ريال ، وذلك باحتمال 0.9 ، 0.1 ، على الترتيب . أحسب توقع مبيعاته اليومية .

١١) في طريقه إلى عمله ، يجتاز موظف ثلاث إشارات ضوئية . والإشارات تعمل مستقلة بعضها عن بعض . واحتمال أن يواجه إشارة حمراء هو 0.4 ، 0.8 ، 0.5 ، بالنسبة للإشارات الثلاث ، على الترتيب .
أ ـ ليكن Y عدد الإشارات الحمر التي يواجهها الشخص في رحلته اليومية إلى عمله ، أوجد توزيع Y .
ب ـ أحسب القيمة المتوقعة لـ Y وانحرافه المعياري .
جـ ـ افترض أن وقت الانتظار لكل إشارة حمراء هو دقيقتان . ما هي القيمة المتوقعة والانحراف المعياري للوقت الذي ينتظره هذا الموظف للرحلة الواحدة .

١٢) لنقم بمحاكاة التجربة في التمرين الثالث بوضع علامات مميزة على ست قطع متماثلة من الورق ، بحيث تمثل اثنتان منها البطاريتين الفاسدتين ، وتمثل القطع

الأربع الباقية البطاريات الصالحة للاستعمال. ضع قطع الورق هذه في قبعة، أخلطها جيدا واسحب ثلاثا منها، ثم سجل قيمة X عدد القطع التي تمثل بطاريات فاسدة. أعد القطع إلى القبعة ثانية وكرر العملية نفسها من جديد حتى تحصل على مائة قياس لـ X . ارسم مدرج التكرار النسبي لهذه العينة من القياسات وقارنه مع المدرج الاحتمالي الذي حصلت عليه من ذلك التمرين.

١٣) في التمرين الثالث أحسب $E(X) = \mu$ ، و$V(X) = \sigma^2$ ، وهما متوسط وتباين X في المجتمع النظري من القياسات، وذلك باستخدام دالة التوزيع التي حصلت عليها هناك. ثم احسب المتوسط $\bar{X}$ ، والتباين S^2 للعينة من القياسات التي حصلت عليها في التمرين ١٢ ، هل يشكل $\bar{X}$ تقديرا جيدا لـ μ ، و S^2 تقديرا جيدا لـ σ^2 ؟

١٤) استخدم المدرج الاحتمالي الذي حصلت عليه في التمرين الثالث لحساب النسبة من مجتمع القياسات الواقعة ضمن انحرافين معيارين على جانبي المتوسط، وقارن مع نظرية تشيبيشيف. أعد في عينة القياسات المذكورة في التمرين ١٢ .

١٥) ولد عينة من 50 قياسا من المجتمع من القياسات الموافق للمتغير X المذكور في المثال (٣ـ٦). وذلك بقذف حجر نرد 50 مرة ، وتسجيل X بعد كل قذفة. احسب $\bar{X}$ و S^2 للعينة، وقارنها مع $E(X)$ و $V(X)$.



الفصل الرابع

نماذج احتمالية لمتغيرات منفصلة

(٤ ـ ١) التجربة الثنائية

يقترن أحد أهم المتغيرات العشوائية المنفصلة بتجربة قذف قطعة نقود، وبالمعنى المجرد للكلمة يُنفذ يوميا العديد من تجارب قذف قطعة النقود ذات الأهمية التطبيقية في العلوم الاجتماعية، والفيزيائية وفي الصناعة وغيرها. . . ، ففي تجارب سبر الرأي العام تشبه مقابلتنا للناخب، من عدة نواح، قذف قطعة نقود. فجوابه «نعم» يوافق وجه الـH، مثلا، وجوابه «لا» أو امتناعه عن الجواب يقابل الحصول على وجه الـ T .

وهناك أمثلة مشابهة في العلوم الاجتماعية، وفي الصناعة، وفي التربية. إذ يهتم الباحث الاجتماعي بنسبة المنازل الريفية المزودة بالكهرباء. وصانع المنظفات يرغب في تقدير نسبة ربات البيوت اللواتي يفضلن نوعا معينا من المنظفات، ويهتم الأستاذ بتقدير نسبة الطلاب الذين سينجحون في مادته. وسنحصل من كل شخص نقابله على ما يشبه نتيجة قذف قطعة نقود «غير متوازنة بصورة عامة».

والرمي في اتجاه هدف معين يشبه قذف قطعة نقود. فإما أن تكون النتيجة إصابة الهدف، أو عدم إصابته. وإطلاق صاروخ إما أن يكون إطلاقا ناجحا أوفاشلا. وإما أن يكون دواء جديد مفيدا، لمعالجة مرض معين أو لا يكون مفيدا. وإذا اخترنا قطعة مصنعة من خط إنتاج صناعي فإما أن تكون خالية من أي عيب صناعي أو تكون معيبة صناعيا. وتكشف مثل هذه التجارب، على تنوعها، ميزات وخواص التجربة الثنائية.

تعريف التجربة الثنائية

التجربة الثنائية هي تجربة تتصف بالخواص التالية :

١ ـ تتألف التجربة من عدد من التكرارات المتماثلة تماما، n مثلا .

٢ ـ يُنتج كل تكرار إحدى نتيجتين، فإما أن تكون النتيجة «نجاحا»، (أي وقوع الأمر الذي نحن في صدد دراسته) وسنرمز لها بـ S، أو أن تكون فشلا، وسنرمز للنتيجة عندئذ بـ F .

٣ ـ احتمال النجاح في تكرار معين، وسنرمز له بـ p يبقى ثابتا من تكرار إلى آخر. ويكون احتمال الفشل، بالطبع، $1 - p$ وسنرمز له بـ q.

٤ ـ التكرارات مستقلة بعضها عن بعض

٥ ـ نهتم بعدد النجاحات التي نحصل عليها خلال التكرارات الـ n ، وسنرمز لهذا العدد بـ X .

وسوف لا تتحقق هذه الشروط جميعها على وجه تام إلا فيما ندر من الحالات العملية. ولكن آثار الحيدان عن هذه الشروط سيبقى بسيطا، ولا يؤثر تأثيرا يُذكر في النتيجة النهائية، طالما بقي هذا الحيدان ضمن حدود معتدلة. فمثلا يبقى احتمال مقابلة ناخب مؤيد للقضية التي ندرسها ثابتا تقريبا من شخص إلى آخر، ما دام مجتمع الناخبين كبيرا جدا بالمقارنة مع العينة من الناخبين الذين تجري مقابلتهم. وإذا كان خمسون بالمائة، مثلا، من مجتمع يحوي ألف ناخب يفضلون المرشح A ، فإن احتمال الحصول على تأييد لـ A في أول مقابلة هو 1/2. واحتمال التأييد في المقابلة الثانية هو 499/999. أو 500/999 ، حسبما تكون المقابلة الأولى قد تمت مع مؤيد أو مع معارض، على الترتيب. والعددان قريبان جدا من 1/2 ، ويمكن اعتبارهما مساويين لـ 1/2 عمليا. كما يمكن الاستمرار في مثل هذا الاعتبار في المقابلة الثالثة والرابعة، وهكذا حتى المقابلة الـ n ، طالما بقي n صغيرا بالنسبة للعدد 1000. وعلى الوجه الآخر، إذا اقتصر المجتمع على عشرة، وكان خمسة منهم يفضلون A، فإن احتمال الحصول على تأييد في المقابلة الأولى هو 1/2 ، ولكنه في الثانية 4/9 أو 5/9، أي أن الاحتمال p يتغير كثيرا من تكرار إلى آخر، ولا يمكن اعتبار التجربة، تجربة ثنائية .

(٤ ـ ٢) دالة التوزيع الثنائي

لنتذكر أولا صيغة نشر ثنائية الحد كما نجدها في كتب الجبر الابتدائية :

$$(q+p)^n = q^n + \binom{n}{1} p\, q^{n-1} + \binom{n}{2} p^2\, q^{n-2} + \dots + \binom{n}{x} p^x\, q^{n-x} + \dots + p^n$$

ونكتب هذا النشر بصورة مختزلة كما يلي :

$$(q+p)^n = \sum_{x=0}^{n} \binom{n}{x} p^x\, q^{n-x}$$

لنتساءل الآن عن دالة توزيع المتغير العشوائي X، وهو عدد النجاحات الملحوظة في تجربة ثنائية خلال n من التكرارات. والمطلوب ببساطة، وكما رأينا في الفصل السابق، هو الاجابة، بصورة عامة، عن السؤال التالي :

ما هو احتمال أن يأخذ المتغير العشوائي X قيمة تساوي x ، أي $P(X = x)$ ؟

وسنجيب عن هذا السؤال في حالة $n = 1$ ، ثم $n = 2$ ، ثم $n = 3$ ، ومن تلمس الخيط المشترك في الحالات الثلاث نحاول استنتاج جواب السؤال المطلوب في الحالة العامة وبالتالي نستنتج صيغة التوزيع .

حالة $n = 1$

الجواب واضح في هذه الحالة من خواص التجربة الثنائية مباشرة، فقيمة X إما أن تكون مساوية للصفر (أي لنتيجة F) أو مساوية للواحد (أي النتيجة S). ونعلم بالفرض أن $P(F) = q$ ، وأن $P(S) = p$. ومنه جدول التوزيع الاحتمالي المطلوب :

جدول (٤ ـ ١) التوزيع الثنائي النقطي أو توزيع برنوللي

X	$f(x)$
0	q
1	p

(ب)

X	الاحتمال	نقاط العينة
0	q	F
1	p	S

(ا)

ويسمى التـوزيع في هذه الحالـة التوزيع الثنـائي النقطي (أو توزيع بيرنـوللي).
ونلاحـظ أن احتمال أن يأخـذ X القيمـة 0 هو الحـد الذي يتضمـن p مرفوعا إلى القوة 0 في عبارة $(p+q)^1 = p+q$. (أي الحد الذي لا يظهـر فيه الحرف p). وأن احتمال أن يأخـذ X القيمة 1 هو الحد الذي يتضمن p مرفوعا إلى القوة 1.

(٤ ـ ٢أ) حالة $n = 2$

في هذه الحالة يبين الجدول (٤ ـ ٢) فضاء العينة والاحتمال الموافق لكل نقطة عينة وقيمة X عند هذه النقطة. ويبين الجدول (٤ ـ ٢) ب دالة التوزيع الاحتمالي لـ X. والقيم الممكنة لـ X هي الآن 0، 1، 2 .

جدول (٤ ـ ٢) التوزيع الثنائي في حالة $n = 2$

نقاط العينة	الاحتمال	X
FF	q^2	0
SF	pq	1
FS	qp	1
SS	p^2	2

(أ)

X	$f(x)$
0	q^2
1	$2pq$
2	p^2

(ب)

وباستعراض العبارة الناتجة عن نشر $(q+p)^2$، وهي

$$q^2 + 2pq + p^2$$

نلاحظ أيضا أن احتمال أن يكون X. مساويا للصفر، أي $f(0)$، هو الحد الذي يحوي p مرفـوعا إلى القـوة صفر (لا تظهـر فيه p)، وأن احتمال أن يكـون X. مساويا للواحـد، أي $f(1)$، هـو الحـد الذي يتضمن p مرفوعا إلى القوة 1، وأن احتمال أن يكون X. مساويا للقيمة 2، أي $f(2)$، هو الحد الذي يتضمن p مرفوعا إلى القوة 2.

(٤ ـ ٢ ب) حالة $n = 3$

يتضمن فضـاء العينة في هـذه الحالة $2^3 = 8$ نقـاط، والقيم الممكنة لـ X (عدد النجـاحـات) هي 0، 1، 2، 3. ويبين الجـدول (٤ ـ ٣)أ و(٤ ـ ٣)ب فضـاء العينـة، ودالة التوزيع، على الترتيب.

جدول (٤ ـ ٣) التوزيع الثنائي في حالة n = 3

نقاط العينة	الاحتمال	X
FFF	q^3	0
SFF	$p\,q^2$	1
FSF	$p\,q^2$	1
FFS	$p\,q^2$	1
FSS	$p^2\,q$	2
SFS	$p^2\,q$	2
SSF	$p^2\,q$	2
SSS	p^3	3

(أ)

X	$f(x)$
0	q^3
1	$3p\,q^2$
2	$3\,p^2\,q$
3	p^3

(ب)

ونلاحظ هنا أيضا انطباق القاعدة التي وجدناها في الحالتين السابقتين. فمن الجدول (٤ ـ ٣) ب نجد أن $f\,(0)$ هو الحد الذي يحوي p مرفوعا إلى القوة صفر في نشر ثنائية الحد

$$(q+p)^3 = q^3 + 3pq^2 + 3p^2q + p^3$$

وأن $f\,(1)$ هو الحد الذي يتضمن p مرفوعا إلى القوة 1، وأن $f\,(2)$ هو الحد الذي يحوي p مرفوعا إلى القوة 2، وأن $f\,(3)$ هو الحد الذي يحوي p مرفوعا إلى القوة 3.

وبتعميم هذه القاعدة نقول، بصورة عامة، أي في حالة n من التكرارات، إن $f\,(x) = P\,(X=x)$ هو الحد الذي يتضمن p مرفوعا إلى القوة x عند نشر ثنائية الحد $(q+p)^n$. ومن صيغة النشر التي استعرضناها في مستهل هذه الفقرة نكتب :

$$f(x) = \binom{n}{x} p^x q^{n-x}\ ;\ x = 0, 1, \ldots, n$$

وهي الصيغة العامة لدالة الاحتمال في حالة التوزيع الثنائي .

*وفيما يلي سنقدم عرضا سريعا لاشتقاق رياضي لهذه الصيغة العامة. فما نبغيه هو الاجابة عن السؤال التالي: ما هو احتمال الحصول على x نجاحا عند تكرار التجربة الثنائية n مرة؟ وللاجابة نقول إن احتمال هذه الحادثة، ولنرمز لها بـB، مثلا، هو مجموع احتمالات نقاط العينة التي تنتمي إلى B. وكل نقطة عينة هي متتابعة من n من الحروف F و S، ولكي تنتمي إلى B يجب أن تحوي الحرف S عددا من المرات يساوي x، وتحوي الحرف F عددا من المرات يساوي $n - x$، أي أن لكل نقطة من B الاحتمال نفسه وهو جداء يحوي x مرة العدد p و $n - x$ مرة العدد q. وهكذا يكون احتمال كل نقطة من نقاط الحادثة B مساويا لـ$p^x q^{n-x}$. ويبقى أن نعرف عدد مثل هذه النقاط التي تتضمنها الحادثة B. ولكن هذا العدد ليس إلا عدد إمكانات تقسيم n موضعا إلى زمرتين، تتضمن إحداهما x موضعا، وتتضمن الأخرى $n - x$ موضعا، وبحيث يظهر الحرف S في مواضع الزمرة الأولى ويظهر في مواضع الزمرة الثانية الحرف F. ونعلم أن هذا العدد هو متوافقات n شيئا مأخوذ x منها في وقت واحد، أي $\binom{n}{x}$. ويكون احتمال الحادثة B هو:

$$f(x) = \binom{n}{x} p^x q^{n-x} \; ; \; x = 0, 1, \ldots, n.$$

وتجدر ملاحظة أن

$$\sum_{x=0}^{n} f(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (q+p)^n = 1^n = 1$$

كما ينبغي أن يكون.

مثال (٤ ـ ١)

لوحظ لفترة طويلة أن صيادا يصيب هدفه باحتمال 0.8. إذا أطلق 4 طلقات على هدف، فما احتمال:

ا ـ إصابة الهدف مرتين؟

ب ـ إصابة الهدف مرتين على الأقل؟

---

* للقراءة فقط.

الحـــل

علينـا أولا تعريف المقصـود بكلمة «نجـاح»، فإذا قلنـا إن النجاح هـو إصابـة الهدف يكون $p = 0.8$، ويصبح $q=1-p=1-0.8=0.2$. والخطوة الثانية هي معرفة عدد التكرارات n، ومن الواضـح هنا أن $n=4$. وبعد تحديد n و p تصبح صيغـة دالة الاحتمال للتوزيع الثنائي محددة تماما. وللإجابـة عن أي سؤال نعبر عنه أولا بدلالة عدد النجاحات X، ثم نطبق صيغة التـوزيع الثنائي لحساب الاحتمال المطلوب. وفي مثـالنا نجد أن صيغة التوزيع الثنائي هي:

$$f(x)=\binom{4}{x}(0.8)^{x}(0.2)^{4-x}\ ;\ x=0,1,2,3,4\ .$$

ا ـ المطلوب $P(X=2)$ أي $f(2)$. وبتعويض x بـ 2 في صيغة التوزيع نجد:

$$f(2)=\binom{4}{2}(0.8)^{2}(0.2)^{2}$$

$$=\frac{4!}{2!\ 2!}(0.64)(0.04)=0.1536\ .$$

ب ـ المطلوب هو $P(X\geq 2)$ ولكن:

$$P(X\geq 2)=P(X=2)+P(X=3)+P(X=4)$$

$$=f(2)+f(3)+f(4)$$

$$=0.1536+\binom{4}{3}(0.8)^{3}(0.2)+\binom{4}{4}(0.8)^{4}$$

$$=0.9728$$

والجدير بـالملاحظة أن هذه الاحتمالات سـوف لا تكون صحيحة إذا قـام الرامي بتفقد موقـع الطلقة في كل مرة. ذلك لأنه سيستفيـد من ملاحظاته في الطلقـة التالية، وعنـدها سيكـون من المتـوقع ازدياد قيمـة p من محاولـة إلى أخـرى، وسـوف لا تكـون التكرارات مستقلة كما يقتضي تعريف التجربة الثنائية.

مثال (٤ ـ ٢)

يجري تفتيش الشحنات الكبيرة من البضاعة القادمة إلى مؤسسة صناعية بطريقة العينة. لنفترض أن هذه الطريقة تتلخص في اختيار عشر قطع عشوائيا، ثم اختبارها واحدة فأخرى. وتُرفض البضاعة إذا لاحظنا قطعتين مرفوضتين أو أكثر.

إذا احتوت شحنة بضاعة على 5% من القطع المرفوضة فما هو احتمال قبول البضاعة؟ رفضها؟

الحــل

إذا عرفنا النجاح بأنه الحصول على قطعة مرفوضة يكون $p = 0.05$, $q = 0.95$. ومن الواضح أن $n = 10$. وصيغة دالة الاحتمال في مثالنا هي :

$$f(x) = \binom{10}{x}(0.05)^{x}(0.95)^{10-x} \; ; \; x = 1, 2, \ldots, 10 .$$

نعبر الآن عن السؤال المطروح بدلالة عدد النجاحات X فنجد :

$$P(\text{قبول البضاعة}) = P(X \leq 1) = P(X \text{ يساوي } 0 \text{ أو } 1) =$$

$$P(X = 0) + P(X = 1) = f(0) + f(1)$$

$$= (0.95)^{10} + \binom{10}{1}(0.05)(0.95)^{9} = 0.914$$

$$P(\text{رفض البضاعة}) = 1 - P(\text{قبول البضاعة}) = 1 - 0.914 = 0.086$$

مثال (٤ ـ ٣)

اختبر لقاح جديد لتحديد فعاليته في الوقاية من الزكام. وقد أعطي لعشرة أشخاص روقبوا لفترة سنة. ووجد أن ثمانية منهم لم يصابوا بالزكام. إذا كان احتمال عدم الإصابة بالزكام خلال سنة هو، بصورة طبيعية، 0.5، فما احتمال ألا يصاب ثمانية أو أكثر علما أن اللقاح لا يزيد في مقاومة الجسم للبرد؟

الحـــل

لنعـرف النجـاح بأنــه عــدم الاصــابة بالزكـام خلال سنـة، فيكون $p = 0.5$، $q = 0.5$، $n = 10$، وتكون دالة الاحتمال لعدد الناجين من الاصابة، X، هي :

$$f(x) = \binom{10}{x} (0.5)\ (0.5)^{10-x}\ ;\ x = 0, 1, 2, \ldots, 10$$

والمطلوب هو $P(X \geq 8)$، ولحسابه نكتب :

$$\begin{aligned} P(X \geq 8) &= P(X = 8) + P(X = 9) + P(X = 10) \\ &= f(8) + f(9) + f(10) \\ &= \binom{10}{8} (0.5)^8 (0.5)^2 + \binom{10}{9} (0.5)^{10} + \binom{10}{10} (0.5)^{10} \\ &= 0.055 \end{aligned}$$

مثال (٤ ـ ٤)

إذا كان 90% من طـلاب مقرر الاحصـاء ينجحـون، فما احتمال فشل اثنين على الأقل من فصل يتضمن عشرين طالبا؟

الحـــل

لنعرف «النجاح» بأنه فشل الطـالب في المقرر. فعندئذ يكون $p = 0.1$ و$q = 0.9$، و$n = 20$. وتكون دالة الاحتمال لعدد الفاشلين، X، هي :

$$f(x) = \binom{20}{x} (0.1)^x (0.9)^{20-x}\ ;\ x = 0, 1, \ldots, 20.$$

أما المطلوب فهو حساب $P(X \geq 2)$. ولدينا :

$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) = 1 - f(0) - f(1) \\ &= 1 - (0.9)^{20} - \binom{20}{1} (0.1)(0.9)^{19} \end{aligned}$$

لاحــظ أن تعـريفنـا للنجاح هنـا كان يتوخى التعبير بسهولـة عن الاحتمال المطلوب. ولو أننـــا عرفنـــا «النجـــاح» بـأنــه نجاح الطالب في المقرر لأصبح $p = 0.9$، $q = 0.1$، $n = 20$، ولأصبحت دالة التوزيع لعدد الناجحين في المقرر، X، هي :

$$f(x) = \binom{20}{x} (0.9)^{x} (0.1)^{20-x} \; ; \; x = 0, 1, \ldots, 20.$$

ويكون المطلوب هو $P(X \le 18)$ لأن فشل اثنين على الأقل يعني أو يكافيء نجاح ثمانية عشر على الأكثر. ولكن

$$P(X \le 18) = 1 - P(X \ge 19) = 1 - P(X = 19) - P(X = 20)$$
$$= 1 - f(19) - f(20) = 1 - \binom{20}{19}(0.9)^{19}(0.1) - \binom{20}{20}(0.9)^{20}$$

وهو الجواب الذي حصلنا عليه سابقا بالضبط.

ونلاحظ من الأمثلة السابقة أن دالة التوزيع الثنائي تقدم علاقة بسيطة لحساب احتمالات حوادث عددية، وهي قابلة للتطبيق في صف واسع من التجارب التي نواجهها في الحياة اليومية. ولكن لابد من الحذر عند استخدام دالة التوزيع الثنائي والتأكد من أن الحالة المدروسة تحقق بصورة مقبولة شروط التجربة الثنائية المذكورة في الفقرة (٤ ـ ١).

وتجدر أيضا ملاحظة أن الأمثلة الأربعة السابقة هي مسائل احتمالية أكثر منها إحصائية. فقد فرضنا أن احتمال النجاح p، وهو الذي يحدد تركيبة المجتمع المدروس، معروف، وكان المطلوب حساب احتمال الحصول على عينة من هذا المجتمع، لها مواصفات محددة. ولو عكسنا الطريقة وافترضنا أننا نملك عينة من مجتمع لا نعرفه ونريد القيام باستقراء حول قيمة p، فعندئذ يقدم المثالان (٤ ـ ٢) و(٤ ـ ٣) مسائل عملية ممتازة يكون الهدف النهائي فيها هو الوصو ل إلى استقراء إحصائي. وسنناقش هاتين المسألتين بتفصيل أكبر في فقرات قادمة.

(٤ ـ ٣) متوسط التوزيع الثنائي وتباينه*

وفقا لتعريف المتوسط والتباين كما ذكرناهما في الفصل السابق نكتب:

$$\mu = E(X) = \sum_{x=0}^{n} x \cdot f(x) = \sum_{x=0}^{n} x \cdot \frac{n!}{x!\,(n-x)!} \; p^{x} q^{n-x}$$

* البراهين الرياضية للقراءة فقط.

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!} \; p^{x} q^{n-x}$$

فقد أغفلنا القيمة 0 للمتغير X لأنها ستؤدي عند تعويضها في الحد العام إلى مقدار يساوي الصفر ($0 \times f(x) = 0$) لنفرض الآن أن $y = x - 1$ فيمكن كتابة العلاقة السابقة على الشكل:

$$\mu = E(X) = \sum_{y=0}^{n-1} \frac{(n-1)!\; n}{y!\,(n-1-y)!} \; p^{y+1} q^{(n-1)-y}$$

$$= np \sum_{y=0}^{n-1} \frac{(n-1)!}{y!\,[(n-1)-y]!} \; p^{y} q^{(n-1)-y}$$

$$= np\,(p+q)^{n-1} = np.$$

وبطريقة مشابهة نحسب $E[X(X-1)]$ فنجد:

$$E[X(X-1)] = \sum_{x=0}^{n} \frac{x(x-1)\; n!}{x(x-1)\,(x-2)!\,(n-x)!} \; p^{x} q^{n-x}$$

$$= \sum_{x=2}^{n} \frac{n(n-1)\,(n-2)!}{(x-2)!\,(n-x)!} \; p^{x} q^{n-x}$$

وبوضع $Y = X - 2$، أي $X = Y + 2$ ، نكتب:

$$E[X(X-1)] = n(n-1)\; p^{2} \sum_{y=0}^{n-2} \frac{(n-2)!}{y!\,(n-2-y)!} \; p^{y} q^{(n-2)-y}$$

$$= n(n-1)\; p^{2}\,(p+q)^{n-2} = n(n-1)\; p^{2}$$

ولكن:

$$E[X(X-1)] = E(X^{2} - X) = E(X^{2}) - E(X)$$

ومنــه:

$$E(\quad) = E[X(X-1)] + E(X) = n(n-1)p^{2} + np$$

وهكذا يكون التباين:

$$\sigma^{2} = E(X-p)^{2} = E(X^{2}) - [E(X)]^{2}$$

$$= n(n-1)p^{2} + np - n^{2}p^{2}$$

$$= np - np^{2} = np(1-p) = npq$$

وهكذا نستنتج القاعدة التالية:

لحساب متوسط التوزيع الثنائي نضرب عدد التكرارات n باحتمال النجاح p.
ولحساب تباين التوزيع الثنائي نضرب عدد التكرارات n باحتمال النجاح p ثم نضرب الناتج باحتمال الفشل q .

مثال (٤ ـ ٥)

قذفنا 400 ربع ريال على منضدة. ما القيمة المتوقعة والانحراف المعياري لعدد القطع التي تُظهر وجه الـH؟

الحــــل

كل ربع ريال يمثل تكرارا لتجربة قذف قطعة نقود (متزنة). وإذا اعتبرنا ظهور وجه الـH نجاحا يكون عدد أوجه الـH الظاهرة، X ، متغيرا يتبع التوزيع الثنائي حيث $n = 400$ و $p = 1/2$. ويكون

القيمة المتوقعة لـ X، $E(X) = np = 400 \times 1/2 = 200$

تباين X، $\sigma_X^2 = npq = 400 \times \frac{1}{2} \times \frac{1}{2} = 100$

والانحراف المعياري لـ X هو الجذر التربيعي للتباين أي

$$\sigma_X = \sqrt{100} = 10$$

وبما أن التكرارات الـn في التوزيع الثنائي ما هي إلا n تكرارا مستقلا لتجربة ثنائية، فيمكن النظر إلى متغير التوزيع الثنائي X على أنه مجموع n من المتغيرات المستقلة X_1، X_2، . . . ، X_n، أي:

$$X = \sum_{i=1}^{n} X_i = X_1 + X_2 + \dots + X_n$$

حيث يخضع كل من هذه المتغيرات في الطرف الأيمن للتوزيع الثنائي النقطي، أي يأخذ كل منها القيمة 1 باحتمال p والقيمة صفر باحتمال $q = 1 - p$. ونعلم من خواص التوقع وخواص التباين أن:

$$V(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

$$V(X) = V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$$

(المتغيرات مستقلة)

لنأخذ الآن X_1 ولنحسب توقعه وتباينه فنجد من الجدول (٤ ـ ٤) ومن تعريف التوقع والتباين:

جدول (٤ ـ ٤) توزيع X_1

x_1	$f(x_1)$
0	q
1	p

$$E(X_1) = 0 \times q + 1 \times p = p$$

$$E(X_1^2) = 0^2 \times q + 1^2 \times p = p$$

$$V(X_1) = p - p^2 = p(1-p) = pq.$$

وكذلك الأمر بالنسبةلـ X_2 وبقية المتغيرات حتى X_n، فتوقع كل منها يساوي p ما دام احتمال النجاح يبقى نفسه من تكرار إلى آخر، وتباين كل منها pq. ومنه يتضح أن:

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np .$$

وأن:

$$V(X) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} pq = npq .$$

وهي النتائج نفسها التي توصلنا إليها بتطبيق مباشر للتعريف.

تمارين (٤ ـ ١)

١) لنفرض أن واحدا من كل عشرة كتب دراسية للمرحلة الجامعية الأولى يصيب نجاحا باهرا. اختارت دار نشر عشرة كتب جديدة لنشرها. فما احتمال:

ا ـ أن ينال واحد منها فقط نجاحا باهرا؟

ب ـ واحد منها على الأقل يصيب نجاحا باهرا؟

جـ ـ اثنان منها على الأقل يصيبان نجاحا باهرا؟

٢) لنفرض أن المحركات الأربعة لطائرة تجارية تعمل مستقلة بعضها عن بعض. وأن احتمال تعطل أي منها والطائرة في الجو هو 0.1 احسب احتمال:

ا ـ ألا يقع أي عطل والطائرة في الجو.

ب ـ ألا يقع أكثر من عطل واحد.

٣) احتمال كشف جهاز رادار لطائرة معادية هو 0.9. إذا كان لدينا خمسة أجهزة، تعمل مستقلة بعضها عن بعض، فاحسب احتمال:

ا ـ ظهور طائرة معادية على شاشات أربعة منها.

ب ـ اكتشاف وجود طائرة معادية في سمائنا.

٤) قذفنا قطعة نقود متوازنة ثلاث مرات. ليكن X عدد أوجه الـH الملحوظة،

ا ـ اكتب دالة الاحتمال لـX، وارسم مدرجها الاحتمالي.

ب ـ احسب متوسط X وانحرافه المعياري.

جـ ـ بالاستفادة من المدرج الاحتمالي، أوجد النسبة من مجتمع القياسات الواقعة ضمن انحراف معياري واحد على جانبي المتوسط، أعد من أجل انحرافين معياريين. هل تتفق نتائجك مع متباينة تشيبيشيف؟

٥) لنفرض أن قطعة النقود غير متوازنة إلى حد كبير وأن احتمال ظهور وجه الـH هو $p = 0.1$. أعد الخطوتين أ و ب في التمرين السابق ولاحظ كيف تفقد دالة الاحتمال تناظرها عندما لا يكون p مساويا للنصف.

٦) ما احتمال أن يكون أربعة على الأقل من أول ستة أشخاص تقابلهم في الشارع في يوم معين قد وُلدوا يوم الجمعة ($7^6 = 117,649$).

٧) إذا أمكن الافتراض أن عدد المواليد الذكور مساو تقريبا لعدد المواليد الاناث في مجتمع سكاني معين. فأوجد النسبة من الأسر ذوات الستة أطفال التي تتصف بما يلي:

ا ـ عدد الأطفال الذكور يساوي عدد الأطفال الاناث .
ب ـ جميع الأطفال الستة من الجنس نفسه .

٨)ارسم المدرج الاحتمالي للتوزيع الثنائي في كل من الحالات التالية :
أ ـ $p = 0.3$ ، $n = 8$ ؛ ب ـ $p = 0.5$ ، $n = 1$ ؛
جـ ـ $p = 0.1$ ، $n=5$؛ د ـ $p = 0.1$ ، $n = 10$.

٩) إذا كان 10% من نوع معين من صمامات التلفزيون يحترق قبل انتهاء مدة الكفالة . وبيع ألف صمام ، فما متوسط وتباين X ، حيث X عدد الصمامات المحترقة قبل انتهاء مدة كفالتها؟ وما الحدود التي تتوقع أن يقع x ضمنها؟ (استخدم متباينة تشبيشيف) .

١٠) يتضمن جانب من امتحان معين 14 سؤالا من النوع متعدد الاختيارات ، وأمام كل سؤال أربعة أجوبة مقترحة ، واحد منها فقط هو الجواب الصحيح .
ا ـ إذا خصص لكل سؤال درجة واحدة فما هي الدرجة المتوقعة لطالب يجيب معتمدا على الحزر (أي يختار جوابه عشوائيا)؟
ب ـ إذا خصص لكل إجابة خاطئة 1- فكم يجب أن نخصص للإجابة الصحيحة حتى تكون الدرجة المتوقعة لطالب يجيب بالحزر صفرا؟

١١) في رحلتك الصباحية إلى الجامعة تضطر إلى اجتياز 12 مجموعة من إشارات المرور تعمل مستقلة بعضها عن بعض . واحتمال أن تكون أي منها خضراء عند وصولك إليها هو 1/2. إذا وقفت عند أقل من 3 مجموعات فستجد وقتا لتناول فنجان من الشاي قبل بداية المحاضرة ، وإذا وقفت عند أكثر من 8 مجموعات فستصل إلى قاعة المحاضرات متأخرا . وإذا تأخرت أكثر من مرتين عن موعد المحاضرة خلال أسبوع يتضمن 5 محاضرات صباحية من السبت إلى الاربعاء ، فستتلقى إنذارا من أستاذك .
ا ـ ما احتمال أن تستطيع تناول فنجان شاي في صباحي السبت والأحد؟
ب ـ ما احتمال أن تتلقى إنذارا من أستاذك؟

١٢) إذا كان 5% من البيض الوارد إلى محل لتسويق المواد الغذائية مكسورا، واشتريت 10صناديق في كل منها 6 بيضات، فما احتمال ألا يحتوي أي منها بيضتين أو أكثر من البيض المكسور؟ ما هو في المتوسط عدد الصناديق الذي سيتضمن بيضتين أو أكثر من البيض المكسور؟

١٣) في مدينة كبيرة كان عدم الوفاق بين الزوجين سببا لـ 60% من حالات الطلاق، أوجد احتمال أن ثلاثا من بين حالات الطلاق الست القادمة في هذه المدينة سُيعزى إلى هذا السبب؟

١٤) إذا كان احتمال أن يحتاج طالب إلى وقت إضافي في اختبار الإحصاء هو 0.1، فأوجد احتمال أن اثنين على الأكثر من 5 طلاب سيحتاجون إلى وقت إضافي. ما احتمال ألا يحتاج واحد على الأقل من الطلاب الخمسة إلى وقت إضافي؟

١٥) إذا كان احتمال تحمل نوع من المصابيح للضغط العالي هو 0.4، وأخذنا عينة من 100 مصباح، فما احتمال ألا يتحمل 65 منها الضغط العالي. (أعط صيغة الجواب دون إجراء الحسابات).

١٦) إذا كان 10% من إنتاج آلة معينة يتضمن عيبا صناعيا، وأخذنا عينة عشوائية من 100قطعة، فما احتمال أن يكون ثلاثون منها على الأكثر معيبا؟ (أعط صيغة الجواب دون إجراء الحسابات).

١٧) من سجلات المواليد في إحدى مدن ولاية يوتا الأمريكية تجد فيما يلي بيانا إحصائيا يُظهر جنس كل من الأطفال الأربعة مرتبة حسب تعاقب ولادتهم في كل من 7745 أسرة من الأسر ذات الأربعة أطفال. (M ترمز لذكر F ترمز لأنثى) والمطلوب:

ا ـ استخدام هذه الإحصائية لتقدير نموذج احتمالي مناسب لمجتمع الأسر من أربعة أطفال. (اعتبر أن التجربة هي أن تختار عشوائيا أسرة من مجتمع الأسر ذات الأربعة أطفال وتسجل جنس الأطفال الأربعة حسب تعاقب ولادتهم).

ب ـ استخـدام هذه الإحصائية لتقدير التوزيع الاحتمالي لـX، عدد الأطفال الذكور في أسرة من أسر هذا المجتمع.

جـــ استخدام هـذه الاحصائية لتقـدير احتمال ولادة طفل ذكـر في هذا المجتمع. واعتباره احتمال النجاح p لتجربة ثنائية فيها $n = 4$ ، و X عدد الذكور من بين الأربعة.

د ـــ اكتب التوزيـــع الاحتمالي لـــX في الســؤال جـ وهـو التوزيع النظـري وقارنـه بالتوزيع الاحتمالي لـX الـذي استنتجته في ب وهـو التوزيع التجـريبي. هل تعتقد أن التوزيع الثنائي هو النمـوذج الاحتمالي المناسب لوصف ودراسة عدد الصبيـان في أسرة من مجتمع الأسر ذوات الأربعـة أطفـال؟ أي وصف مجتمع القياسات للمتغير العشوائي X، الذي يرمز إلى عـدد الذكور في أسرة من هذا المجتمع؟ (قم بحساباتك لثلاثة أرقام عشرية).

التكرار	جنس الأطفال في الأسرة حسب ترتيب ولادتهم	التكرار	جنس الأطفال في الأسرة حسب ترتيب ولادتهم
526	MFFM	537	MMMM
498	FMFM	549	MMMF
490	FFMM	514	MMFM
429	MFFF	523	MFMM
451	FMFF	467	FMMM
456	FFMF	497	MMFF
441	FFFM	486	MFMF
408	FFFF	473	FMMF

١٨) وجدت شركة طيران أنه، في المتوسط، يفشل 4 بالمائة من المسافرين الذين يحجزون مقاعد لرحلة معينة في الوصول إلى قاعة المسافرين في الوقت المحدد. ولذلك قررت

الشركة السماح لـ 75 شخصا أن يحجزوا مقاعدهم في طائرة لا تتسع إلا لثلاثة وسبعين راكبا. ما احتمال توفر مقعد لكل مسافر يصل في الوقت المحدد؟

(٤ ـ ٤) الكشف على بضاعة بطريقة العينة*

نعلم أن المؤسسة الصناعية هي مكان تتحول فيه مادة أو مواد خام إلى مادة مصنعة. ولابد لإدارة المؤسسة، حفاظا منها على مستوى معين لجودة المصنوعات، أن تجعل كمية المادة الخام غير الصالحة التي تدخل في عملية الإنتاج أصغر ما يمكن. كما ترغب في خفض عدد القطع المنتجة المعيبة صناعيا إلى أقل حد ممكن أيضا. وفي محاولة لبلوغ هذا الهدف تقيم «غربالا» للبضائع الداخلة في عملية الإنتاج والخارجة منها في محاولة لمنع غير المناسب من العبور في الاتجاهين كليهما.

ولتبسيط المناقشة، لنفرض أن ما يهمنا هو «غربلة» البضاعة الواردة، أي المواد الخام المؤلفة، مثلا، من قطع على شكل صناديق من مادة معينة. فإما أن نقبل شحنة البضاعة الواصلة إلى المصنع، إذا كانت نسبة غير الصالح فيها نسبة مقبولة. وإما أن تكون هذه النسبة عالية فنرفض البضاعة ونردها إلى المموّل.

ويمكن إقامة «الغربال» بعدة طرق. ومن الواضح أن أكمل هذه الطرق هي أن يتم الكشف على كامل البضاعة قطعة فأخرى. وللأسف فإن تكاليف مثل هذا الكشف قد تكون كبيرة إلى الحد الذي يجعلها غير واردة البتة من وجهة النظر الاقتصادية. هذا ناهيك عما يمكن أن يُقبل أو يُرفض خطأ من قبل المفتش، خاصة بعد أن ينال منه الجهد مناله نظرا لضخامة العمل المطلوب. يضاف إلى ذلك أنه قد تكون هذه الطريقة مرفوضة بالنظر إلى طبيعة المادة التي نكشف عليها. فاختبار صلاحية المصباح الصغير (الفلاش) المستخدم في آلة تصوير يؤدي إلى تلفه. واختبار كل البضاعة، في مثل هذه الحالة، يعني ألا يبقى شيء لاستخدامه أو لبيعه.

وطريقة الغربلة الثانية الأقل تكلفة، والتي توفر جهودا كبيرة، هي طريقة العينة الإحصائية. وهي مشابهة للخطة التي ذكرناها في المثال (٤ ـ ٢). وفيها نختار عينة من

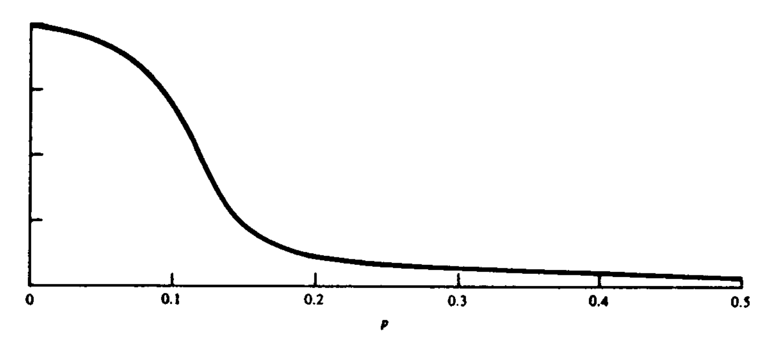
* للقراءة فقط.

n قطعة من قطع الشحنة بطريقة عشوائية، ونكشف عليها بدقة قطعة فأخرى لمعرفة ما تحويه العينة من قطع غير صالحة. وإذا كان عدد هذه القطع، ولنرمز له بـ X، أقل أو يساوي عددا a حددناه سلفا، ويسمى عدد القبول، نقرر قبول البضاعة، وفيما عدا ذلك نرفضها ونعيدها إلى المموّل. وكان عدد القبول في الخطة التي ذكرناها في المثال (٤ ـ ٢) هو $a=1$.

ويلاحظ القارىء أن خطة العينة تعمل بطريقة موضوعية تماما، وتؤدي إلى استقراء يتعلق بمجتمع القطع التي تتألف منها الشحنة. ورفض البضاعة يعني أننا استقرأنا أن نسبة القطع غير المقبولة، p، هي نسبة كبيرة تتجاوز الحد الذي يمكن التساهل فيه، والذي يؤدي إلى تدهور مستوى الجودة في الناتج النهائي في المصنع. وقبول البضاعة يعني أننا استقرأنا أن تلك النسبة، p، صغيرة، وأنها تبقى في حدود المعقول في عملية التصنيع. ويقدم الكشف على بضاعة بطريقة العينة مثالا على عملية اتخاذ قرار إحصائي.

ولا تكون مناقشتنا تامة إذا أهملنا بعض النقاط المتعلقة بجودة الطريقة المستخدمة للقيام بالاستقراء. ومع أن خطة العينة التي عرضناها أعلاه هي طريقة لاتخاذ قرار إلا أنها ليست وحيدة. ويمكننا تغيير حجم العينة n، عدد القبول a، أو اتباع طريقة في اتخاذ القرار غير إحصائية وراجعة للتقديرات الشخصية. فكيف يمكن مقارنة هذه الطر ق المختلفة في اتخاذ القرار؟ والجواب الطبيعي هو أن نختار الطريقة التي تؤدي إلى القرار الصحيح بأكبر تواتر ممكن، أو على الوجه الآخر تؤدي إلى القرار غير الصحيح بأقل نسبة من المرات.

ويميز مهندسو الإنتاج جودة خطة العينة بحساب احتمالات قبول البضاعة في حالة نسب مختلفة للقطع غير الصالحة في الشحنة الواردة. ويمثلون نتائج هذه الحسابات في شكل بياني يدعى «المنحنى العملياتي المميز» لخطة العينة. ويبين الشكل (٤ ـ ١) نموذجا لمثل هذه المنحنيات. ولكي يؤدي الغربال مهمته بصورة مرضية، نرغب في أن يكون احتمال قبول شحنات نسبة العطل فيها ضعيفة مرتفعا، وأن يكون منخفضا في حالة شحنات نسبة العطل فيها مرتفعة. ويلاحظ القارىء أن احتمال القبول سينحدر باستمرار مع ارتفاع نسبة العطل، وهي النتيجة التي نتوقعها.



شكل (٤ ـ ١) نموذج لمنحنى عملياتي مميز لخطة عينة .

وعلى سبيل المثـال ، إذا كان الممـوّل مطمئنا إلى أن نسبـة العطل في شحناتـه من البضاعة لا تتجاوز 1% ، وكان المصنع يعمل بصورة مرضيـة بشحنات تقل نسبة العطل فيها عن 5%، فعندئذ يجب أن يكون احتمال قبول شحنات بنسبة من العطل أقل من 1% مرتفعا . ومـا لم يكن الأمر كذلك فإن المموّل سيرفع أسعاره لتغطية نفقـات إعادة شحنة ممتازة «تحوي أقل من 1% من العطل» إليـه ، أو أنه سيحمّل إدارة المصنع نفقـات إعادة الكشف عـلى البضـاعـة . وعلى الوجـه الآخر فإن احتمال قبول شحنـات نسبة العطل فيها 5% أو أكثر لابد أن يكون منخفضا .

مثال (٤ ـ ٦)

احسـب احـتمال قبول شحنـة عند استخـدام خطة عينـة فيهـا حجـم العينة $n = 5$ ، وعـدد القبـول $a = 0$. وذلـك إذا كانـت نسب القطـع غيـر الصالحـة تساوي $p = 0.1$، $p = 0.3$، $p = 0.5$. ارسم المنحنى العمليـاتي المميـز لهذه الخطة .

الحـل

لدينا توزيع ثنائي فيه $n = 5$ ، واحتمال النجاح p . وصيغة دالة التوزيع :

$$f(x) = \binom{5}{x} \; p^x \; q^{5-x} \; , \; x = 0, 1, 2, 3, 4, 5$$

ويكون

$$P(\text{القبول}) = f(0) = q^5$$

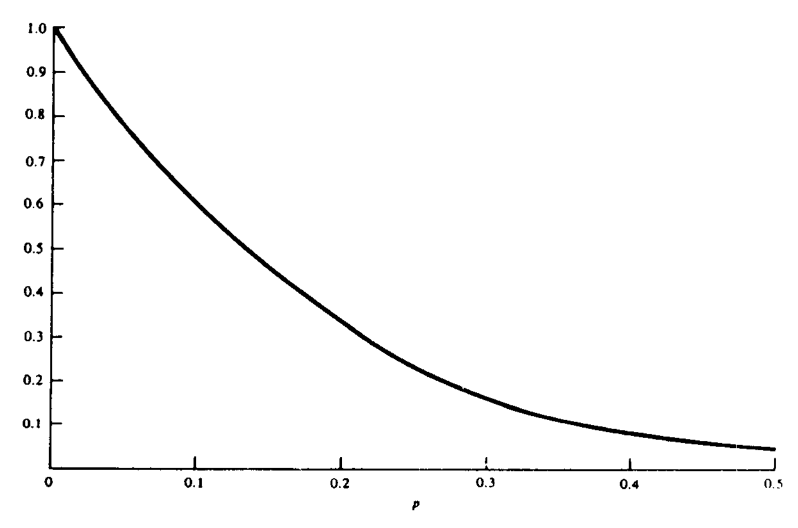
ومنه:

$$P(\text{القبول} \mid p = 0.1) = (0.9)^5 = 0.590$$

$$P(\text{القبول} \mid p = 0.3) = (0.7)^5 = 0.160$$

$$P(\text{القبول} \mid p = 0.5) = (0.5)^5 = 0.031$$

ونعلم بالاضافة إلى ذلك أن احتمال القبول يجب أن يكون الواحد عندما يكون $p = 0$، وأن يكون صفرا عندما $p = 1$. وبرسم النقاط الخمسة، حيث الإحداثي السيني هو نسبة العطل. والإحداثي الصادي هو احتمال القبول الموافق، يمكن تخطيط شكل تقريبي للمنحنى العملياتي المميز وهو مبين في الشكل (٤ ـ ٢).



شكل (٤ ـ ٢) لمنحنى العملياتي المميز في حالة $n = 5$ ، $a = 0$

وقد يصبح حساب احتمالات التوزيع الثنائي عملا شاقا في حالة n كبيرة. ولتيسير الحسابات تتوفر عادة جداول تعطي مجموع احتمالات التوزيع الثنائي من $x = 0$ إلى $x = a$ = عدد القبول. وذلك في حالة عينات حجمها n يساوي 5، 10، 15، 20، 25 .

مثال (٤ ـ ٧)

ارسم المنحنى العملياتي المميز لخطة عينة فيها $n = 15$ و $a = 1$.

الحـــل

سنحسب احتمال القبول في حالة $p = 0.1$ ، $p = 0.2$ ، $p = 0.3$ ، $p = 0.5$. وهكذا نكتب :

$$\text{احتمال القبول} = \sum_{x=0}^{1} f(x) = f(0) + f(1) = q^{15} + \binom{15}{1} p\, q^{14}$$

ومنه :

$$P(\text{القبول} \mid p = 0.1) = (0.9)^{15} + 15\ (.1)(.9)^{14} = .549$$

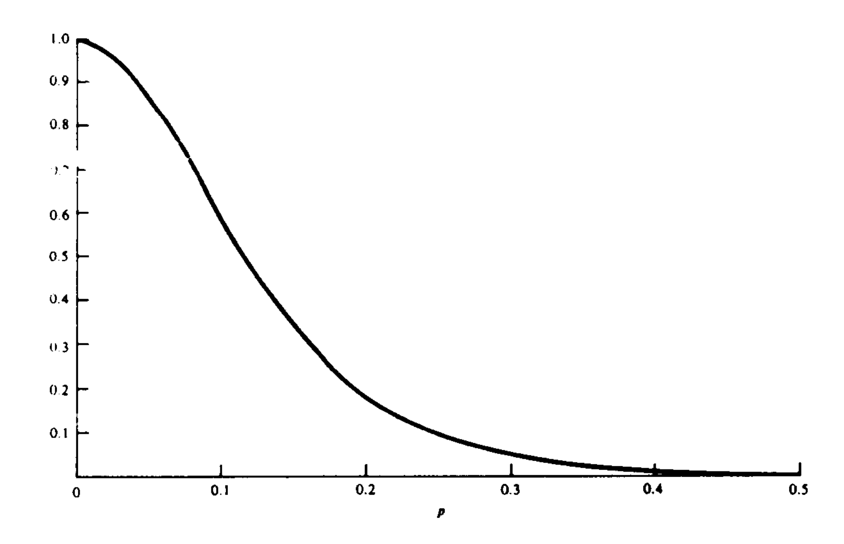
وبصورة مشابهة نجد :

$$P(\text{القبول} \mid p = 0.2) = 0.167$$

$$P(\text{القبول} \mid p = 0.3) = 0.035$$

$$P(\text{القبول} \mid p = 0.5) = 0.000$$

والمنحنى العملياتي المميز مبين في الشكل (٤ ـ ٣) .



شكل (٤ ـ ٣) : المنحنى العملياتي المميز في حالة $n = 15$ ، $a = 1$

وتستخدم خطة العينة على نطاق واسع في الصناعة . ولكل خطة عينة منحن عملياتي مميز، يميز الخطة عن غيرها ، ويقدم نوعا من الوصف لحجم ثقوب الغربال . وسيختار مهندس الإنتاج الخطة بحيث يحقق المتطلبات التي يفرضها واقعه . فزيادة

عدد القبـول تزيد من احتمال القبـول، وبالتالي توسع ثقـوب الغربال . كما تقـدم زيادة حجم العينة قدرا أكبر من المعلومات التي يمكن أن نبني عليها قرارنا، وبالتالي تزيد من قدرة الطـريقة المتبعـة على التمييز. و هكذا ينحـدر المنحنى العملياتي المميـز بسرعة مع ازدياد p عنـدما يكون حجم العينة n كبيرا. (قـارن بين الشكـل (٤ ـ ٢) حيث $n = 5$ ، والشكل (٤ ـ ٣) حيث $n = 15$) .

تمارين (٤ ـ ٢)

١) يتفق شـار وبـائع على استخدام طريقـة الكشف بـالعينة مستخدمين عينة حجمها $n = 5$ وعدد قبول $a = 0$. ما هو احتمال أن يقبل الشاري شحنة بضاعة نسبة العطل الحقيقية فيها:
أ ـ $p = 0.1$، ب ـ $p = 0.3$ ، جـ ـ $p = 0.5$ ، د ـ $p = 0$، هـ ـ $p = 1$.
ارسم المنحنى العملياتي المميز لهذه الخطة.

٢) أعد التمرين ١ في حالة $n = 5$ ، $a = 1$.

٣) أعد التمرين ١ في حالة $n = 10$ ، $a = 0$.

٤) أعد التمرين ١ في حالة $n = 10$ ، $a = 1$.

٥) ارسم المنحنيات العمليـاتية المميـزة للخطط الأربع في التمارين ١ ، ٢ ، ٣، ٤، على ورقة بيانيـة واحدة. ما تأثير زيـادة عدد القبول a مع بقاء n ثابتة؟ ومـا تأثير زيادة حجم العينة n، عندما تبقى a ثابتة؟

(٤ ـ ٥) اختبار فرضية*

إن مسألة اللقـاح ضد الزكـام المعطاة في المثـال (٤ ـ ٣) ، هي مسألة تـوضيحية لاختبار إحصـائي لفرضيـة. وتتلخص المسألة في السؤال التـالي: هل تقدم المعلـومات التي تحويها العينة دلالة كافية على فعالية اللقاح؟

* للقراءة فقط.

ويحمل المنطق المستخدم في اختبار فرضية شبهـا كبيرا بالأسلوب المستخدم في قاعة محكمة. فعند محاكمة رجل متهم تفترض المحكمة أن المتهم بريء حتى تثبت إدانته. ويجمع ممثل النيابة كل الأدلة المتوفرة له ويقدمها في محاولة لنقض فرضية البراءة، وبالتالي الحصول على إدانة المتهم والحكم عليه. وتصور المسألة الاحصائية اللقاح ضد الزكام متهما. والفرضية التي سيجري اختبارها، وتدعى الفرضية الابتدائية، هي أن اللقاح غير فعال. ودلالات الدعوى موجودة ضمن العينة المسحوبة من مجتمع. ويعتقد الباحث وهو يؤدي دور ممثل النيابة أن اللقاح مفيد فعلا. ويحاول تبعا لذلك استخدام الدلالات المتوافرة في العينة لرفض الفرضية الابتدائية وبالتالي دعم قناعته بأن اللقاح، في الحقيقة، ناجح جدا ضد الزكام. (لاحظ أن التهمة هنا هي أن اللقاح فعال، والفرضية الإبتدائية هي براءة اللقاح من هذه التهمة) وسيتعرف القارىء على هذا الأسلوب كشكل أساسي من أشكال الطريقة العلمية الحديثة حيث يتوجب وضع النظريات المقترحة على محك الواقع.

ويبدو بديهيا أن نختار عدد من يجانبهم الزكام، X، كقياس لمقدار البينة التي تحتويها العينة. وإذا كان X كبيرا فإننا سنميل إلى رفض الفرضية الإبتدائية واستنتاج أن اللقاح فعال. وعلى الوجه الآخر سيقدم صغر X القليل من الدعم لرفض الفرضية الإبتدائية. وفي الحقيقة، إذا كانت الفرضية الإبتدائية صحيحة واللقاح غير فعال فإن احتمال النجاة من الزكام طوال فصل الشتاء سيكون $p = 1/2$، وستكون القيمة المتوسطة لـX هي:

$$E(X) = n\,p = 10\left(\frac{1}{2}\right) = 5\,.$$

وقد لا يجد معظم المهتمين صعوبة في تكوين حكمهم الخاص في حالة $X = 10$ أو في حالة يساوي 5، 4، 3، 2، أو 1 حيث تقدم في الظاهر دلالة كبيرة لرفض أو قبول الفرضية، على الترتيب. ولكن ماذا يمكن أن يُقال في حالات أقل وضوحا مثل $X = 7$ أو $X = 8$، أو $X = 9$؟ وسواء اتخذنا قرارا بطريقة ذاتية أو موضوعية، فمن الواضح أننا سنختار الطريقة التي تعطي أقل احتمال لاتخاذ قرار غير صحيح.

وسيختبر الإحصائي الفرضية الابتدائية بطريقة موضوعية، ولكنها مشابهة لما يمكن أن نصل إليه باللجوء إلى الحس السليم أو الفطرة. وصانع القرار، ويدعى عادة «الإحصاء» يُحسب عادة من العينة. وفي مسألتنا فإن هذا الإحصاء هو عدد من نجوا من الإصابة بالزكام، X، وسنأخذ عندئذ في اعتبارنا كل القيم الممكنة لهذا الإحصاء، وهي هنا، $X = 0, 1, 2, \ldots, 10$ ثم نقسم هذه القيم إلى مجموعتين، ندعو إحداها منطقة الرفض، والأخرى منطقة القبول. وهكذا تُنفذ التجربة، ونلاحظ قيمة «صانع القرار» أو «الإحصاء»، X. فإذا أخذ X قيمة من منطقة الرفض رفضنا الفرضية. وفيما عدا ذلك نقبلها. وعلى سبيل المثال، يمكننا اختيار منطقة الرفض من النقاط $X = 8$ أو 9، أو 10. ونعتبر ما تبقى من قيم X منطقة قبول. وبما أننا لاحظنا القيمة 8 في مثالنا فإننا نرفض الفرضية الإبتدائية بأن اللقاح غير فعال ونستنتج أن احتمال النجاة من الزكام طوال عام كامل هي أكبر من 1/2 عند استخدام اللقاح. والآن ما هو احتمال أن نرفض الفر ضية الإبتدائية مع أنها في الواقع صحيحة؟ واحتمال الرفض الخاطىء للفرضية الإبتدائية هو احتمال أن نأخذ X القيمة 8، 9، أو 10، علما أن $p = 1/2$، وهذا هو، في الحقيقة، الاحتمال الذي حسبناه في المثال (٤ ـ ٣) ووجدناه مساويا لـ 0.055 . وبما أننا قررنا رفض الفرضية الإبتدائية ووجدنا أن احتمال أن يكون هذا الرفض غير صحيح هو احتمال بسيط فإن هذا يولد لدينا ثقة غير قليلة بأننا اتخذنا القرار الصحيح.

عند تأمل المسألة قليلاً سيلاحظ القارىء أن الشركة المنتجة للقاح ستواجه نوعين من الخطأ. فمن جهة يمكن أن ترفض الفرضية الإبتدائية وتستنتج خطأ أن اللقاح فعال. وإنتاج الدفعة الأولى من اللقاح وطرحها للاستخدام سيسبب خسارة مادية ومعنوية (الإساءة إلى سمعة الشركة) لأن الحقيقة ستكشف عن نفسها. ومن الجهة الأخرى، يمكن أن تقرر قبول الفرضية الابتدائية، وتستنتج خطأ أن اللقاح غير فعال. وسيقود هذا الخطأ إلى خسارة الفوائد الجمة الصحية والمادية التي كان سيقدمها طرح ذلك اللقاح المفيد في الأسواق لاستخدامه على نطاق واسع.

ويدعى رفض الفرضية الابتدائية مع أنها صحيحة بالخطأ من النوع الأول (أو النوع I). ونرمز لاحتمال ارتكاب خطأ من النوع الأول بـ α . وسيزداد الاحتمال α أو

يتناقص مع اتساع أو تقلص منطقة الـرفض. وبالقدر الذي تمثل فيه α مخاطرة الرفض الخاطىء، يمكن أن نتسـاءل: لماذا لا نختـار منطقـة الرفض صغيرة قـدر المستطـاع، ونقلل بذلك احتمال تلك المخاطـرة؟ فمثلا لماذا لا نختار $X = 10$ فقط منطقة رفض في مثالنا هنا؟ ولكن لسوء الحظ إن تخفيض α يزيد من احتمال ارتكاب خطأ من نوع آخر، وهو احتمال قبـول الفرضيـة الابتدائية مع أنها غير صحيحـة، وأن الصحيح هو فـرضية بديلـة تختلف عنها. ويـدعى هذا الخطأ بـالخطأ من النوع الثـاني (أو النوع II). ونرمز لاحتمال مثل هذا الخطأ بـالرمز β. أي أن β هـو احتمال القبـول الخاطىء. ومن أجل حجم ثابت للعينة n، تكون العلاقة بين α و β علاقة عكسية. فعنـدما يزداد أحدهما يتناقص الآخر. وتقدم زيادة حجم العينة معلـومات أكثر يمكن أن نبني عليها قرارنا، وبالتالي تخفض كلا من α و β. ويقيس احتمالا الخطأ من النـوعين I و II، أي α و β، مخاطـرة التـورط بقـرار غير صحيح. ويختـار المجـرب، وفقـا لما تمليـه طبيعـة المسألـة المدروسة، حجم هذين الاحتمالين. وعادة نختار حجم العينة n، ونحدد شكل منطقة الـرفض وحجمهـا، بحيث نضع سقفـا لـلاحتمال α لا يتجـاوزه، ويسمـى مستـوى المعنوية، وتحت هـذا الشرط نحاول جعل β أصغر ما يمكن. ومـن الواضح أن اختيار شكل منطقة الرفض يشكل أمـرا حاسما في مسألة الاختبار الإحصـائي وتتوقف عليه إلى حد كبير قوة وكفاءة الاختبار الإحصائي.

تمارين (٤ ـ ٣)

١) نقوم بتجربة لاختبـار أن قطعة نقود متوازنة، وذلك بقذف قطعـة النقود أربع مرات وملاحظة عدد أوجه الصورة التي تظهر. ونـرفض الفرضية إذا كان هذا العدد صفرا أو أربعة.

ا ـ ما هو احتمال الخطأ من النوع الثاني؟

ب ـ إذا كانت القطعة فعلا غير متوازنة واحتمال ظهور وجه الصورة هو 0.7، فما هو احتمال الخطأ من النوع الثاني في هذا الاختبار؟

٢) نتـوقع أن يعطي زوج من الخنافس نسـلا بعينين سوداوين بنسبـة 30% من المرات. ولاختبار هـذه النظرية نـلاحظ ثلاثا من نسلهـا فنجد أن عيونها زرقـاء. فهل تقدم هذه النتيجة دلالة كافية لنقض النظرية؟ علل إجابتك إحصائيا.

٣) نفّذنا عددا من تجارب علم النفس كما يلي: استدرجنـا فأرا إلى نهاية حاجز يتفرع منه ممران يقود كـل منهما إلى باب. وهدف التجربة أسـاسا هو تحديد مـا إذا كان للفأر قدرة على تفضيـل أحد الممريـن. في تجربة مؤلفـة من 6 محاولات لوحظـت النتائج التالية:

المحاولة	1	2	3	4	5	6
الباب الذي اختير	2	1	2	2	2	2

أ ـ عبر عن الفرضية التي تود اختبارها.
ب ـ ليكن X عدد المرات التي يختار الفأر فيهـا الباب الثاني، فما هي قيمة α في هذا الاختبار إذا احتوت منطقة الرفض على $X = 0$ و $X = 6$؟
جـ ـ ما هي قيمة β من أجل الفرضية البديلة $p = 0.8$؟

٤) سجلنـا عدد المآخـذ الكهربـائية التـي تحوي عيبا صنـاعيا في كل من خطي إنتـاج مختلفين A و B، وذلك يوميا ولمدة عشرة أيام فحصلنا على النتائج التالية:

اليوم	1	2	3	4	5	6	7	8	9	10
الخط A	172	165	206	184	174	142	190	169	161	200
الخط B	201	179	159	192	177	170	182	179	169	210

إذا علمت أن حجم الإنتاج الكلي هـو نفسه بـالنسبة للخطين. قـارن عدد القطع المعيبة الناتجة عن الخطين كل يوم، وليكن X عدد الأيام التي يكون فيها B متجاوزا لـ A، فهل تقدم هـذه المعلومات دلالـة كافيـة للقول بأن الخط B ينتج في المتـوسط قطعا معيبة أكثر من A؟ إعرض الفرضيـة التي ستختبرها واستخدم X إحصاء لهذا الاختبار.

(٤ ـ ٦) توزيع بواسون*

لتوزيع بواسون مجالات تطبيق واسعة، فهو يقدم، على وجه العموم، نموذجا جيدا للمعلومات الاحصائية التي تأخذ شكل تعداد لحوادث نادرة الوقوع. ويمثل المتغير العشوائي البواسوني، X مثلا، عدد «الحوادث النادرة» الملحوظة في وحدة قياس معينة، زمنا كانت أم مسافة أم مساحة أم حجما. ونوضح بالأمثلة التالية، التي نطبق فيها عادة توزيع بواسون:

١ ـ ليكن X عدد المكالمات الهاتفية، في شركة معينة، كل خمس دقائق من الفترة الممتدة بين الساعة الثانية عشرة ظهرا والساعة الثانية بعد الظهر.
٢ ـ ليكن X عدد قوالب الزبدة المباعة خلال يوم في محلات بيع المواد الغذائية.
٣ ـ ليكن X عدد الأعطال الأسبوعية الناشئة عن العجلات في أسطول من شاحنات النقل البري.
٤ ـ ليكن X عدد الجسيمات الصادرة في الثانية عن كمية من مادة مشعة.
٥ ـ ليكن X عدد الأخطاء المطبعية في الصفحة عبر صفحات كتاب معين.
٦ ـ ليكن X عدد الالكترونات التي يُصدرها مهبط مسخن في فترة زمنية محددة.
٧ ـ ليكن X عدد ذرات الغاز في منطقة جزئية صغيرة v من وعاء مليء بهذا الغاز حجمه V.
٨ ـ ليكن X عدد حوادث السيارات في مدينة كبيرة خلال يوم.
٩ ـ ليكن X عدد البكتريا الموجودة في مم٣ من وعاء يحتوي على سائل معين.

وتكفي هذه الأمثلة لتوضيح مدى تنوع واتساع تطبيقات التوزيع البواسوني.

(٤ ـ ٦ ـ ١) دالة الاحتمال لتوزيع بواسون*

يمكن استنتاج دالة الاحتمال لتوزيع بواسون كحالة حدية (أو كنهاية) لدالة الاحتمال للتوزيع الثنائي. فلنفرض أن عدد التكرارات n يسعى في اتجاه أن يصبح كبيرا

---

* للقراءة فقط.

جدا وأن p يسعى في اتجاه أن يصبح صغيرا جدا، وبحيث يبقى جداؤهما $n\,p$ مساويا لعدد ثابت λ، مثلا، ولنكتب دالة الاحتمال للتوزيع الثنائي على الشكل:

$$f(x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!\,(n-x)!}\; p^x \frac{(1-p)^n}{(1-p)^x}$$

$$= n(n-1)\ldots(n-x+1)\frac{\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^n}{x!\left(1-\frac{\lambda}{n}\right)^x}$$

$$= \frac{n}{n}\;\frac{n-1}{n}\;\ldots\;\frac{n-x+1}{n}\;\frac{\lambda^x\left(1-\frac{\lambda}{n}\right)^n}{x!\left(1-\frac{\lambda}{n}\right)^x}$$

ومن أجل قيمة كبيرة جدا لـ n، وقيمة لـ x ثابتة وصغيرة بالمقارنة مع n، تكون كل من النسب $\frac{n-1}{n}$ ، . . . ، $\frac{n-x+1}{n}$ و $\left(1-\frac{\lambda}{n}\right)^x$ قريبة جدا من الواحد، ويمكن كتابة $f(x)$ بصورة تقريبية على الشكل:

$$f(x) \doteq \frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n$$

حيث $\doteq$ تعني يساوي تقريبا. وإحدى نتائج التحليل الرياضي المعروفة هي أن المقدار $\left(1+\frac{1}{n}\right)^n$ يسعى إلى عدد ثابت $e = 2.7183...$ عندما تسعى n إلى اللانهاية. وأن $\left(1-\frac{\lambda}{n}\right)^n$ يسعى إلى العدد $e^{-\lambda}$ وهكذا تصبح الصيغة الحدية لعبارة $f(x)$ كما يلي:

$$f(x) = e^{-\lambda}\,\frac{\lambda^x}{x!}\;;\; x = 0, 1, 2, \ldots$$

وهي صيغة دالة الاحتمال لتوزيع بواسون. وستعطي هذه الصيغة احتمالات مساوية تقريبا لتلك التي تعطيها دالة الاحتمال للتوزيع الثنائي شريطة أن يكون n كبيرا جدا ويكون الجداء $n\,p$ صغيرا نسبيا (نطلب عادة أن يكون $n\,p < 5$).

ويبرهن أن كلا من متوسط توزيع بواسون وتباينه يساوي λ. أي أن

$$E(X) = V(X) = \lambda$$

ويمكن اعتبار هذه الخاصة كخاصة مميزة ينفرد بها التوزيع البواسوني من بين التوزيعات المنفصلة جميعها. وإذا وجدنا في مجتمع من القياسات أن متوسطه وتباينه قريبان جدا من بعضهما فإن ذلك يحفزنا على الاعتقاد بأن أفضل نموذج احتمالي مناسب لهذا المجتمع قد يكون النموذج البواسوني.

مثال (٤ ـ ٨)

يتلقى عامل الهاتف في شركة معينة المكالمات الهاتفية بمعدل مكالمتين في الدقيقة.

أ ـ ما احتمال ألا يتلقى العامل أية مكالمة خلال فترة دقيقة؟

ب ـ ما احتمال وصول مكالمتين خلال فترة دقيقة واحدة؟

جـ ـ ما احتمال ألا يتلقى أية مكالمة خلال فترة خمس دقائق؟

الحـــل

لتحديد دالة الاحتمال لتوزيع بواسون تكفي معرفة λ، وهو يمثل متوسط عدد الحوادث التي تقع في وحدة قياس معينة، هي في مثالنا هنا وحدة قياس زمني وتساوي دقيقة واحدة. إذا $\lambda = 2$، على أساس أن وحدة قياس الزمن هي الدقيقة. ولكن كم تصبح λ لو أن وحدة قياس الزمن أصبحت 5 دقائق بدلا من دقيقة واحدة؟ والجواب واضح، لأنه إذا كان متوسط عدد المكالمات يساوي 2 لكل دقيقة فهو يساوي 10 لكل خمس دقائق. وتجدر ملاحظة أن X في توزيع بواسون يمثل عدد الحوادث التي تقع في وحدة قياس. و $f(x) = P(X = x)$ هو احتمال أن تقع x حادثة في وحدة قياس. وهذا يشير إلى ضرورة التعرف على وحدة القياس أو تحديدها ومن ثم حساب λ وكتابة صيغة دالة الاحتمال الموافقة. وبعد ذلك التعبير عن السؤال المطلوب بدلالة x وحساب الاحتمال المطلوب.

وفي مثالنا وحدة القياس هي الدقيقة بالنسبة للسؤالين أ و ب. وتكون λ كما ذكرنا مساوية لـ 2، فنكتب دالة الاحتمال كما يلي:

$$f(x) = e^{-2}\frac{2^x}{x!}\ ;\ x = 0, 1, 2, \ldots$$

أ ـ المطلوب هو $P(X=0)$ أي $f(0)$ وبتعويض x بصفر في الدالة $f(x)$ نجد:

$$f(0) = e^{-2}\frac{2^0}{0!} = e^{-2} = 0.135 .$$

ب ـ المطلوب هو $P(X=2)$ ، أي $f(2)$. وبتعويض x بـ 2 في $f(x)$ نجد:

$$f(x) = e^{-2}\frac{2^2}{2!} = 2e^{-2} = 0.270$$

جـ ـ باعتبار وحدة القياس الزمني الآن هي «خمس دقائق»، تصبح $\lambda = 10$ وتصبح دالة الاحتمال كما يلي:

$$f(x) = e^{-10}\frac{10^x}{x!} \; ; \; x = 0, 1, 2, \ldots$$

والمطلوب هو $P(X=0) = f(0)$. وبتعويض x بصفر في هذه الدالة نجد:

$$f(0) = e^{-10}\frac{10^0}{0!} = e^{-10} = 0.000045$$

مثال (٤ ـ ٩)

قام بيتمان بتطبيق توزيع بواسون في مسألة فيزيائية مهمة. فقد استخدم دالة الاحتمال البواسونية في تفسير بيان إحصائي تجريبي كان قد جمعه عالمان عظيمان من رواد الفيزياء الذرية هما رذرفورد وجايجر. فقد قاما بتعداد جسيمات α التي انبعثت عن قرص مطلي بالبولونيوم وذلك خلال فترة زمنية تساوي 7.5 ثانية. وسجلا مشاهداتهما في 2608 فترات زمنية متلاحقة. وكان مجموع عدد الجسيمات الملحوظة يساوي 097 10 جسيما. أي أن متوسط عدد الجسيمات الصادرة هو 3.87 جسيما لكل 7.5 ثانية. (أي لكل وحدة قياس حيث وحدة القياس هنا هي 7.5 ثانية). وقد بين بيتمان أنه إذا كانت نظريتهما الذرية صحيحة وكان λ متوسط عدد الجسيمات الصادرة خلال فترة زمنية محددة، فإن X عدد الجسيمات الصادرة خلال فترة زمنية هو متغير عشوائي يخضع لتوزيع بواسون بوسيط (أو معلمة) يساوي λ ، وهكذا إذا استخدمنا 3.87 كأفضل قيمة تخمينية لـ λ متوفرة لنا، فإن نظريتهما الذرية تتنبأ بأن X متغير عشوائي بواسوني دالة احتماله هي:

$$f(x) = e^{-3.87}\frac{(3.87)^x}{x!} \; ; \; x = 0, 1, 2, \ldots$$

ويبين الجدول (٤ ـ ٥) النتائج الواقعية والنتائج النظرية، والتوافق الملحوظ القائم بين المشاهدات التجريبية والتنبؤات النظرية.

جدول (٤ ـ ٥) القيم الواقعية والقيم النظرية لتجربة رذرفوردوجايجر

	عدد الفترات الزمنية (7.5 ثا) التي صدر خلالها n جسيم α	
n	القيمة الملحوظة	القيمة النظرية (مع تدوير الرقم العشري)
0	57	54
1	203	210
2	383	407
3	525	525
4	532	508
5	408	394
6	273	254
7	139	140
8	45	68
9	27	29
10	10	11
11	4	4
12	0	1
≥13	2	1

مثال (٤ ـ ١٠)

تقع حوادث اصطدام الطرق في منطقة معينة بمعدل حادث واحد لكل يومين.

ا ـ احسب الاحتمالات الموافقة لـ 0، 1، 2، 3، 4، 5، 6 حوادث اصطدام في الأسبوع في تلك المنطقة.

ب ـ ما عدد الاصطدامات الأسبوعية الأكثر احتمالاً؟

جـ ـ كم يوما في الأسبوع تتوقع أن يمر بدون اصطدامات؟

الحل

ا ـ متوسط عدد الحوادث لكل أسبوع هو $\lambda = 7\,(0.5) = 3.5$. ودالة الاحتمال هي:

$$f(x) = e^{-3.5} \frac{(3.5)^x}{x!} \quad ; x = 0, 1, 2, \ldots$$

ومنه:

$$f(0) = e^{-3.5} \frac{(3.5)^0}{0!} = 0.030 \;;\; f(1) = e^{-3.5} \frac{3.5}{1!} = 0.106$$

$$f(2) = e^{-3.5} \frac{(3.5)^2}{2!} = 0.185 \;;\; f(3) = e^{-3.5} \frac{(3.5)^3}{3!} = 0.216$$

$$f(4) = e^{-3.5} \frac{(3.5)^4}{4!} = 0.189 \;;\; f(5) = e^{-3.5} \frac{(3.5)^5}{5!} = 0.132$$

$$f(6) = e^{-3.5} \frac{(3.5)^6}{6!} = 0.077 \,.$$

ب ـ من السؤال أ نلاحظ أن عدد الحوادث الأسبوعية الأكثر احتمالا هو $x = 3$.

جـ ـ باعتبار اليوم هو وحدة القياس بدلا من الأسبوع نجد أن $\lambda = 0.5$ وتكون دالة الاحتمال

$$f(x) = e^{-0.5} \frac{(0.5)^x}{x!} \quad ; x = 0, 1, \ldots$$

حيث x الآن هو عدد الحوادث اليومية. ومرور يوم بدون حوادث يعني أن $x = 0$. ولحساب $P(X = 0)$ نعوض x بصفر في دالة الاحتمال فنجد:

$$f(0) = e^{-0.5} = 0.607$$

ونحن الآن أمام مسألة توزيع ثنائي. لنعّرف النجاح بأنه مرور يوم بدون حوادث، باحتمال النجاح $p = 0.607$، في كل يوم من أيام الأسبوع السبعة. وبأخذ $n = 7$، يكون عدد النجاحات هو عدد أيام الأسبوع التي تمر بدون حوادث. إذا رمزنا لهذا العدد بـ y، فإن y متغير عشوائي يخضع للتوزيع الثنائي حيث $n = 7$ و $p = 0.607$. والمطلوب هو $E(y)$. وكما نعلم فإن:

$$E(y) = n\,p = 7 \times 0.607 = 4.25$$

وبصورة تقريبية نقول إنه لو أحصينا عدد أيام الأسبوع التي تمر بدون حوادث في تلك المنطقة وذلك لعدد هائل من الأسابيع، ثم حسبنا متوسط الأعداد التي حصلنا عليها لكان الناتج 4.25 يوما.

تمارين (٤ ـ ٤)

١) تتلقى تحويلة للهاتف المكالمات بين الساعة العاشرة صباحا والثانية عشرة ظهرا بمعدل مكالمتين في الدقيقة. ما هو احتمال ألا تتلقى التحويلة أية مكالمة خلال دقيقة؟ أن تتلقى مكالمتين خلال دقيقة؟ أن تتلقى مكالمتين خلال خمس دقائق؟ ألا تتلقى أية مكالمة خلال خمس دقائق؟

٢) لنفرض أن مساحة صغيرة من زجاجة مجهرية لفحص الدم تحتوي في حالة شخص طبيعي على عشر كريات حمر في المتوسط. ما احتمال أن تتضمن زجاجة من دم شخص طبيعي، في تلك المساحة الصغيرة، أقل من 6 كريات حمر؟ ألا تحوي أية كرية حمراء؟

٣) تتضمن صحيفة يومية في المتوسط ثلاثة أخطاء مطبعية للصفحة الواحدة. ما احتمال أن:

ا ـ تكون الصفحة الأولى خالية من الأخطاء المطبعية؟

ب ـ يوجد ستة أخطاء مطبعية في الصفحة الأخيرة؟

جـ ـ يوجد أكثر من ثلاثة أخطاء مطبعية في صفحة الرياضة.

٤) يبيع مخزن نوعا معينا من الأجهزة الكهربائية بمعدل أربعة في الأسبوع. بافتراض أن عدد الأجهزة المباعة أسبوعيا متغير بواسوني، أوجد عدد الأجهزة التي يجب توافرها في مستودع المخزن في بداية أسبوع بحيث يطمئن صاحب المخزن باحتمال 95% إلى أنه سيلبي جميع الطلبات من هذا النوع من الأجهزة خلال ذلك الأسبوع.

٥) تصل السيارات إلى مرآب في وسط المدينة بمعدل سيارة كل دقيقة. وسيسبب وصول أكثر من أربع سيارات في أي دقيقة أزمة في حركة المرور. كم أزمة تتوقع، في المتوسط، في ساعات العمل الـ 12 في اليوم؟

٦) من بين 150 مباراة في كرة القدم جرت يوم الخميس لم تُسجل أية أهداف في 12 منها . مفترضا توزيع بواسون ، كم تعتقد أن يكون متوسط عدد الأهداف للمباراة الواحدة؟

احسب احتمالات أن :

ا ـ يسجل أقل من هدفين في مباراة معينة .

ب ـ يسجل أكثر من هدفين ولكن أقل من 5 أهداف في مباراة معينة .

٧) تمر المركبات من نقطة معينة على طريق مزدحم بمعدل 300 مركبة في الساعة . أوجد احتمال ألا تمر أي مركبة خلال دقيقة معينة . ما العدد المتوقع للمركبات التي تمر خلال دقيقتين . أوجد احتمال أن يمر بالفعل هذا العدد المتوقع خلال أي فترة طولها دقيقتان؟

٨) سجل عدد حوادث الاصطدام في منطقة معينة يوميا ولفترة امتدت 1500 يوم ، وكانت النتائج كما يلي :

عدد الاصطدامات في اليوم	0	1	2	3	4	5
التكـــرار	342	483	388	176	111	0

ما التوزيع النظري الذي يمكن استخدامه نموذجا مناسبا لهذا البيان؟ احسب التكرارات المتوقعة مستخدما التوزيع الاحتمالي النظري بمتوسط يساوي متوسط البيان الإحصائي أعلاه .

٩) في مسح كبير تناول أكثر من 100 000 ولادة تبين أن معدل الاصابة بمرض في العمود الفقري هي 4.12 لكل ألف ولادة . ما احتمال أن نلاحظ في عينة عشوائية من خمسين ولادة :

ا ـ عدم وجود إصابات؟

ب ـ إصابة واحدة؟
جـ ـ إصابيتين
د ـ أكثر من إصابتين؟
(لاحظ من طريقة اشتقاق التوزيع البواسوني أنه عندما تكون n غير صغيرة، هنا $n = 50$، ويكون احتمال النجاح، صغيرا جدا، هنا $p = .00412$، فيمكن اعتبار التوزيع البواسوني بمتوسط $\lambda = np$ تقريبا جيدا للتوزيع الثنائي.)

(٤ ـ ٧) العينة العشوائية

قلنا إن هدف الإحصاء كعلم هو القيام باستقراء حول خصائص مجتمع اعتمادا على المعلومات التي تحويها عينة مأخوذة من هذا المجتمع. وكل مسألة إحصائية تبدأ بعينة من القياسات أو المشاهدات. وعلى سبيل المثال، عند اتخاذ قرار برفض أو قبول شحنة بضاعة واردة إلى مصنع، وكذلك عند اختبار فرضية تتعلق بفعالية لقاح جديد ضد الزكام، اعتمدنا، في كل حالة، على عينة مأخوذة من مجتمع، ووصفنا العينة بأنها عشوائية، فماذا نبتغي من وصف العينة الاحصائية بأنها عينة عشوائية؟ وفي المقام الأول، متى نقول إن العينة عشوائية؟

ولقد أوضحنا ، في مسألة اختبار فرضية، أنه لابد من حساب احتمال الحصول على عينة كالعينة التي بين أيدينا. (العينة التي تمخضت عنها التجربة) فإذا وجدنا أنها من النوع غير المحتمل (احتمال الحصول عليها تحت الفرضية الابتدائية هو احتمال زهيد) نستنتج أن الفرضية الابتدائية غير مبررة، ونرفضها. وإذا وجدنا أن العينة محتملة تماما قلنا إن الدلالات المتوفرة من العينة لا تسمح لنا برفض الفرضية ولذلك نقبلها. والنقطة المهمة التي نريد إبرازها هي أنه لابد لنا من حساب احتمال الحصول على عينة كتلك التي لاحظناها كي نصل إلى استقراء إحصائي، أو نتخذ قرارا إحصائيا. ولا يخفى أن لطريقة أخذ العينة أثرا حاسما في حساب احتمالها. ومن هنا تأتي أهمية كون العينة عشوائية. فالعشوائية هي الخاصية التي ستجعل حساب مثل ذلك الاحتمال ممكنا، لا بل ستجعله سهلا وميسورا. ولكن متى نقول إن العينة عشوائية؟

لنفرض الآن أننا سحبنا عينة تتضمن n قياسا من مجتمع يحوي N قياسا، فما هو عدد العينات المختلفة التي يمكن الحصول عليها؟ إن هذا العدد، كما نعلم، هو عدد متوافقات N شيئا مأخوذ n منها في وقت واحد، أي

$$\binom{N}{n} = \frac{N!}{n!\,(N-n)!}$$

وفي الفصل الثالث علقنا على مسألة الاختيار العشوائي لعنصر من مجموعة تتضمن N عنصرا، فقلنا إن عشوائية الاختيار تعني أن لكل من العناصر الـ N، الفرصة نفسها في أن يكون العنصر الذي يقع عليه الاختيار. أي أن احتمال الاختيار هو $1/N$ لكل عنصر من العناصر الـ N في المجموعة التي نختار منها. وهذا تعبير كمي عن طريقة اختيار نطمئن فيها إلى عدم إمكانية وجود أي شكل من أشكال التحيز لعنصر دون آخر. وسنطبق الفكرة نفسها لتعريف عشوائية العينة.

تعريف العينة العشوائية

عند سحب عينة حجمها n من مجتمع منته يتضمن N عنصرا، نقول إن العينة عشوائية، إذا كان لكل من العينات الـ $\binom{N}{n}$ الممكنة الفرصة نفسها في أن تكون العينة الملحوظة. أي إذا كان احتمال الحصول على أي منها هو $\frac{1}{\binom{N}{n}}$

وفي معظم التطبيقات الاحصائية يكون المجتمع لانهائيا (أي أن عدد عناصره غير محدود)، وتجريدا ذهنيا أكثر منه عناصر محسوسة. لنأخذ، مثلا، حالة القيام بقياس ثابت فيزيائي في تجربة مخبرية، ولنفرض أننا كررنا التجربة نفسها عشر مرات، فعندئذ ننظر إلى القياسات العشرة الناتجة على أنها عينة من مجتمع افتراضي هو ذلك المجتمع من القياسات التي كنا سنحصل عليها لو أننا قمنا، وبصورة مستقلة، بتكرار التجربة نفسها مرة بعد أخرى إلى ما لانهاية له. وكل قياس من القياسات العشرة هو في الواقع متغير عشوائي قائم بذاته، ويوافقه بالطبع مجتمع من القياسات. وعشوائية العينة تضمن لنا أن المجتمعات العشرة من القياسات، هي، في الحقيقة، مجتمع واحد، وفوق ذلك تضمن لنا أن هذه المتغيرات العشرة تتحول، أو تعمل، مستقلة بعضها عن بعض. وبعبارة مبسطة نقول إنه كي نحصل على عينة عشوائية من n قياسا، ما علينا

إلا أن نكرر التجربة، تحت نفس الشروط والظروف، n مرة. وبطريقة تسمح لنا بالقول إن هذه التكرارات الـn مستقلة فيما بينها، أي لا يمكن أن يكون لنتيجة أي تكرار منها أثر سلبي أو إيجابي على ما يمكن أن تكونه نتيجة تكرار آخر.

ولو أمعنا النظر فيما نقوله وتذكرنا، على سبيل المقارنة، ما قلناه عند تعريف تجربة ثنائية، لوجدنا أن التكرارات الـ n لتجربة ثنائية إنما تمثل تماما عينة عشوائية حجمها n من مجتمع القياسات الموافق لمتغير ثنائي نقطي (متغير بيرنوييللي). فثبات قيمة p من تكرار إلى آخر يلخص شرط ثبات ظروف التجربة، وأننا نكرر التجربة نفسها مرة بعد أخرى، وهو الشرط الأول من شرطي عشوائية العينة، أما استقلال التكرارات بعضها عن بعض فيستكمل الشرط الثاني المطلوب. وسنستخدم هذه الملاحظة المهمة للوصول إلى طريقة حسابات تقريبية، إلا أنها سهلة ومفيدة، في الفصل القادم. وتسمى «طريقة تقريب التوزيع الثنائي بالتوزيع الطبيعي».

(٤ ـ ٨) المعاينة بدون إرجاع والتوزيع فوق الهندسي

عند سحب عينة عشوائية من مجتمع منته (يتضمن عددا محدودا من العناصر)، إذا سحبنا العنصر وسجلنا نتيجة السحب ثم أعدنا العنصر إلى المجتمع قبل سحب عنصر آخر، سُميت المعاينة «معاينة مع الإعادة». أما إذا احتفظنا بالعنصر المسحوب وقمنا بالسحب التالي من العناصر الباقية في المجتمع، أي لم نقم بإعادة العنصر المسحوب إلى المجتمع، سُميت المعاينة «معاينة بدون إعادة». وإذا تضمن المجتمع N عنصرا، مثلا، فسيبقى المجتمع على حاله، بدون تغيير، في الحالة الأولى، إذ يجري، على الدوام، سحب عنصر من بين N عنصرا، هي مجمل عناصر المجتمع. ومن الواضح أن نتيجة كل سحب ستكون مستقلة عن نتيجة أي سحب آخر. وعند سحب عينة حجمها n تكون عمليات السحب الـn تكرارات مستقلة للتجربة نفسها، مما يتفق تماما مع شروط التوزيع الثنائي. أما إذا كانت المعاينة بدون إعادة فإن نتيجة كل سحب ستتأثر بنتائج جميع عمليات السحب السابقة. ولا تشكل عمليات السحب الـ n تكرارات مستقلة بعضها عن بعض، ولا ينطبق عليها بالتالي التوزيع الثنائي. وإذا

كان الأثر زهيدا، n صغيرة جدا بالنسبة لـ N، أي إذا كان الحيدان عن شروط التوزيع الثنائي في حدود طفيفة، فيمكن تطبيق التوزيع الثنائي كتقريب جيد. وفيما عدا ذلك لابد لنا من التفكير في توزيع يتلاءم وشروط المعاينة. وسنجد أن التوزيع فوق الهندسي هو التوزيع الملائم لمعاينة بدون إعادة فما هو التوزيع فوق الهندسي، وما هي الحالات التي نلجأ فيها إلى هذا التوزيع؟

لنفرض أن مجتمعا يتضمن N عنصرا من بينها N_1 عنصرا يتصف بصفة معينة، A مثلا، والعناصر الباقية وعددها $N - N_1$ لا تتصف بالصفة A. إذا سحبنا من هذا المجتمع، وبدون إعادة، عينة عشوائية حجمها n، وعرفنا المتغير العشوائي X بأنه عدد عناصر العينة التي تتصف بالصفة A. فالتوزيع الاحتمالي لـ X يسمى التوزيع فوق الهندسي. وللوصول إلى صيغة هذا التوزيع $f(x)$، علينا أولا تحديد القيم الممكنة لـ X، ثم الاجابة على السؤال التالي:

ما احتمال أن يكون X مساويا لقيمة ما x، حيث ترمز x لأي قيمة من القيم الممكنة. أي ما هو احتمال الحادثة $X = x$، ونكتب:

$$f(x) = P(X = x)$$

ومن الواضح أن القيم الممكنة لـ X هي:

$$0, 1, 2, \ldots, \min(N_1, n)$$

حيث يعني الرمز $\min(N_1, n)$ أصغر العددين N_1, n. فالقيمة x لا يمكن أن تتجاوز n باعتبارها تمثل جزءا من العينة، ولا يمكنها، على الوجه الآخر، أن تتجاوز N_1 لأن العينة، وهي جزء من المجتمع، لا يمكن أن تتضمن من عناصر الصفة A أكثر مما في المجتمع من هذه العناصر.

ولحساب $f(x)$، نتذكر أن التجربة هي سحب عينة عشوائية بدون إعادة حجمها n من المجتمع الموصوف أعلاه، ثم تسجيل x عدد العناصر في هذه العينة التي تتصف بالصفة A. وفضاء العينة هو مجموعة النتائج الممكنة، أي مجموعة كل العينات الممكنة وعددها $\binom{N}{n}$. وبما أن العينة عشوائية فلكل منها الاحتمال نفسه وهو $1 / \binom{N}{n}$. أي

أننا هنا في حـالة نموذج الاحتمالات المتساويـة . ويكون احتمال الحادثة $X = x$ هو عدد الحالات الملائمـة ، . أي عدد العينـات التي تتضمن xمن عناصر الصفة A ، مقسوما على عدد الحالات الممكنـة وهو$\binom{N}{n}$. ولكن عدد الحالات الملائمـة هو عدد إمكـانات اختيار x من N_1 و $n - x$ من $N - N_1$ ، أي $\binom{N_1}{x}\binom{N-N_1}{n-x}$ وهكذا نجد :

$$f(x) = \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, 2, \dots, \min(N_1, n).$$

وهي صيغة التوزيع فوق الهندسي .

ويمكن البرهان على أن متوسط هذا التوزيع :

$$E(X) = n\frac{N_1}{N}.$$

أي n مضروبا بنسبة تواجد عناصر الصفة A في المجتمع . كما يمكن ا البرهان على أن تباين التوزيع $V(X)$ هو:

$$V(X) = \frac{N-n}{N-1}\ n\frac{N_1}{N}\left(1 - \frac{N_1}{N}\right)$$

مثال (٤ ـ ١١)

خضع اثنـا عشر مريضـا بالتهاب القصبـات إلى تجربة طبيـة فاختير ستـة منهم بصورة عشوائية وأعطوا المعالجة A بينما أعطي الباقون المعالجة B. بكم طريقة يمكن توزيع الاثني عشر مريضا على المعالجتين؟ وإذا كان أربعة من المرضى يعانون من ارتفاع ضغط الدم والآخرون طبيعيون بالنسبة لمعدل ضغط الدم ، فما احتمالات :

أ ـ أن يخضع ذوو الضغط المرتفع جميعهم للمعالجة A؟

ب ـ أن يخضع ذوو الضغط المرتفع جميعهم للمعالجة نفسها؟

جـ ـ أن يوجد واحد على الأقل من ذوي الضغط المرتفع في كل معالجة؟

الحـــل

عدد إمكانات توزيع المرضى على المعالجتين هو $\binom{12}{6}$ لأن اختيار 6 وتخصيصهم للمعالجة A يعني بطبيعة الحال تخصيص الستة الباقين للمعالجة B.

$$\binom{12}{6} = \frac{12!}{6!\ 6!} = \frac{12.\ 11.\ 10.\ 9.\ 8.\ 7}{6.\ 5.\ 4.\ 3.\ 2.\ 1} = 924$$

أ ـ لدينا هنا $N = 12$ ، $N_1 = 4$ ، $N - N_1 = 8$ ، $n = 6$ ، وليكن X عدد المصابين بالضغط المرتفع الذين يجري تخصيصهم للمعالجة A. فالتوزيع الاحتمالي لـX هو التوزيع فوق الهندسي.

$$f(x) = \frac{\binom{4}{x}\binom{8}{6-x}}{\binom{12}{6}}\ ;\ x = 0, 1, 2, 3, 4\ .$$

والمطلوب في أ هو احتمال $x = 4$ و هو $f(4)$. وهكذا نكتب:

$$P(X=4) = f(4) = \frac{\binom{4}{4}\binom{8}{2}}{\binom{12}{6}} = \frac{1}{33}$$

ب ـ السؤال هنا يعني أن يكون $x = 4$، جميع ذوي الضغط المرتفع خاضعون للمعالجة A أو $x = 0$، جميع ذوي الضغط المرتفع خاضعون للمعالجة B. وهكذا نكتب:

$$P(X = 4\ \text{أو}\ 0) = P(X=0) + P(X=4) = f(0) + f(4)$$

$$= \frac{\binom{4}{0}\binom{8}{6}}{\binom{12}{6}} + f(4) = \frac{2}{32}$$

جـ ـ هذه الحادثة هي الحادثة المتممة للحادثة المذكورة في ب واحتمالها يساوي

$$1 - \frac{2}{32} = \frac{31}{33}\ .$$

أو إذا حسبنا المطلوب بصورة مباشرة نجده :

$$P(3 \text{ أو } 2 \text{ أو } 1) = f(1) + f(2) + f(3) = \frac{31}{33}$$

مثال (٤ ـ ١٢)

عجنة تتضمن 3 عناصر من النوع أ و 5 عناصر من النوع ب. سحبنا عينة عشوائية* من أربعة عناصر من هذه العجنة. ليكن X عدد العناصر في العينة من النوع أ.

١ ـ اكتب صيغة التوزيع الاحتمالي لـ X، وجدولا يتضمن كل قيمة من القيم الممكنة لـ X، والاحتمال المقابل لها.

٢ ـ احسب $E(X)$ و $V(X)$ باستخدام الجدول (أي بتطبيق التعريف مباشرة). ثم باستخدام الصيغتين المعطاتين لمتوسط التوزيع وتباينه وقارن النتيجتين.

٣ ـ إذا كان عمر كل عنصر من النوع أ سنتين، وعمر كل عنصر من النوع ب خمس سنوات، فما هي القيمة المتوقعة لعمر العينة؟ وما هو تباين عمر العينة؟

الحل

١ ـ يتضمن المجتمع ثمانية عناصر، أي $N = 8$ منها $N_1 = 3$ عناصر من النوع أ. إذا سحبنا عينة عشوائية من أربعة عناصر، $n = 4$، فيكون التوزيع الاحتمالي لـ X هو، بوضوح، التوزيع فوق الهندسي:

$$f(x) = \frac{\binom{3}{x}\binom{5}{4-x}}{\binom{8}{4}}, \quad x = 0, 1, 2, 3.$$

والجدول المطلوب هو:

x	0	1	2	3
$f(x)$	$\frac{1}{14}$	$\frac{6}{14}$	$\frac{6}{14}$	$\frac{1}{14}$

* عبارة «سحبنا عينة عشوائية» ستعني دائما معاينة بدون إعادة ما لم يُذكر غير ذلك.

٢ ـ باستخدام الجدول :

$$E(X) = 0 \times \frac{1}{14} + 1 \times \frac{6}{14} + 2 \times \frac{6}{14} + 3 \times \frac{1}{14} = \frac{3}{2}$$

$$E(X^2) = 0^2 \times \frac{1}{14} + 1^2 \times \frac{6}{14} + 2^2 \times \frac{6}{14} + 3^2 \times \frac{1}{14} = \frac{39}{14}$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{39}{14} - \frac{9}{4} = \frac{156 - 126}{56} = \frac{15}{28} .$$

وعلى الوجه الآخر، كان يمكن التعويض في صيغتي متوسط التوزيع وتباينه لنجد :

$$E(X) = n\frac{N_1}{N} = 4 \times \frac{3}{8} = \frac{12}{8} = \frac{3}{2} ,$$

$$V(X) = \frac{N-n}{N-1}\, n\frac{N_1}{N}\left(1 - \frac{N_1}{N}\right)$$

$$= \frac{8-4}{8-1} \times 4 \times \frac{3}{8}\left(1 - \frac{3}{8}\right)$$

$$= \frac{4}{7} \times \frac{12}{8} \times \frac{5}{8} = \frac{15}{28}$$

وهو بالضبط ما وجدناه بالاستخدام المباشر للتعاريف .

٣ ـ لنرمز لعمر العينة بـ Y فيكون :

$$Y = 2X + 5(4 - X) = 20 - 3X$$

والمطلوب $E(Y)$ و $V(Y)$. ولكن نعلم من خواص التوقع وخواص التباين أن :

$$E(Y) = E(20 - 3X) = 20 - 3E(X)$$

$$= 20 - 3 \times \frac{3}{2} = \frac{31}{2} .$$

$$V(Y) = V(20 - 3X) = 9V(X) = 9 \times \frac{15}{28} = \frac{135}{28} .$$

(٤ ـ ٩) توزيع $\bar{x}$ متوسط عينة من مجتمع منته

نحتاج في عمليات الاستقراء الإحصائي، أشد ما نحتاج، إلى التوزيعات الاحتمالية لخصائص العينة. ما كان منها مقياسا للنزعة المركزية، أو ما كان منها مقياسا للتشتت. وفي الطليعة منها جميعا نجد متوسط العينة $\bar{X}$.

لنفترض أن لدينا مجتمعا منتهيا فيه N عنصرا. إذا سحبنا من هذا المجتمع عينة عشوائية حجمها n ، ورمزنا لمتوسطها بـ $\bar{X}$ ، فماذا نقصد بتوزيع $\bar{x}$ ؟ التوزيع الاحتمالي هو، عمليا، وصف وتحديد لبنية مجتمع القياسات الموافق لمتغير عشوائي، والمتغير العشوائي، كما نعلم، ينبغي أن يكون معرفا على فضاء عينة، وفضاء العينة هو مجموعة النتائج الممكنة لتجربة، فأين التجربة هنا وما فضاء العينة؟ من الواضح أن التجربة هي سحب عينة عشوائية حجمها n من مجتمع يتضمن N عنصرا، وبالتالي ليس فضاء العينة هنا إلا مجموعة كل العينات التي يمكن الحصول عليها. ويجدر الانتباه إذا إلى أن ما يؤخذ في الاعتبار ليس عينة واحدة، سحبناها وحسبنا متوسطها، ولكن مجمل العينات التي كان يمكن أن نحصل عليها لو أننا كررنا تجربة السحب مرة بعد أخرى. وهنا نضع اليد من جديد على الطبيعة التكرارية للمسألة الاحتمالية والمسألة الاحصائية. صحيح أننا نعتمد على المعلومات التي تقدمها العينة للقيام باستقراء حول المجتمع الذي جاءت منه. ولكننا لا نعتمد على هذه المعلومات كقطعة معزولة قائمة بذاتها، وإنما نعتمد عليها، في سياق شريط متكامل، أو جزء من صورة متكاملة، تتضمن العينة التي بين أيدينا وغيرها من العينات التي كنا سنحصل عليها لو أننا كررنا أخذ عينة ثانية وثالثة وهلمجرا. والتوزيع الاحتمالي للعينة أو، على وجه التحديد، لخاصة من خصائصها، ولنقُل $\bar{X}$ مثلا، هو الذي يقدم وصفا لمحتويات تلك الصورة المتكاملة. فمثلا، ما احتمال أن تكون قيمة $\bar{X}$ أكبر من عدد معين؟ أو بعبارة عملية، ما نسبة العينات التي يزيد متوسطها على عدد معين؟ وذلك من بين كل العينات الممكنة؟ ومن خلال التوزيع الاحتمالي لـ $\bar{X}$ يمكننا الإجابة على أسئلة كهذه، كما يمكننا، بصورة عامة، الحكم بأن العينة التي حصلنا عليها هي، مثلا من النوع غير

المحتمل، أو من النوع المحتمل، الأمر الذي ساعدنا عند مناقشة مسألة اختبار فرضية على اتخاذ موقف إحصائي من الفرضية، وكان له الدور الأساسي في بلورة مثل ذلك الموقف.

وللوصول إلى التوزيع الاحتمالي للمتغير العشوائي $\bar{X}$ نحسب قيمته عند كل نقطة عينة أي لكل عينة من العينات الـ$\binom{N}{n}$ الممكنة. ويكون الاحتمال الموافق لكل من القيم المختلفة لـ $\bar{X}$ هو $1/\binom{N}{n}$ مضروبا بعدد المرات الذي تكرر فيه ظهور تلك القيمة. وسنوضح الطريقة بمثال.

مثال (٤ ـ ١٣)

لدينا المجتمع من الأرقام 1,2,3,4,5,6. اكتب التوزيع الاحتمالي لـ $\bar{X}$ متوسط عينة حجمها 2 نسحبها عشوائيا من هذا المجتمع. وارسم المدرج الاحتمالي لهذا التوزيع.

الحـــل

عدد نقاط العينة، أي عدد كل العينات الممكنة هو $15 = \binom{6}{2}$ والاحتمال الموافق لكل منها هو 1/15، وذلك وفقا لتعريف العينة العشوائية. والجدول (٤ ـ ٦) يبين العينات المختلفة الممكنة والاحتمال الموافق لكل منها، والقيمة التي يأخذها المتغير العشوائي $\bar{X}$ في كل نقطة عينة، أي قيمة المتوسط الحسابي للعينة.

ومن هذا الجدول يمكننا بسهولة وضع جدول التوزيع الاحتمالي المطلوب، وهو يتضمن القيم المختلفة لـ $\bar{X}$ والاحتمال الموافق لكل منها. ونلاحظ أن القيم المختلفة لـ $\bar{X}$ هي

$$1.5,\ 2,\ 2.5,\ 3,\ 3.5,\ 4,\ 4.5,\ 5,\ 5.5$$

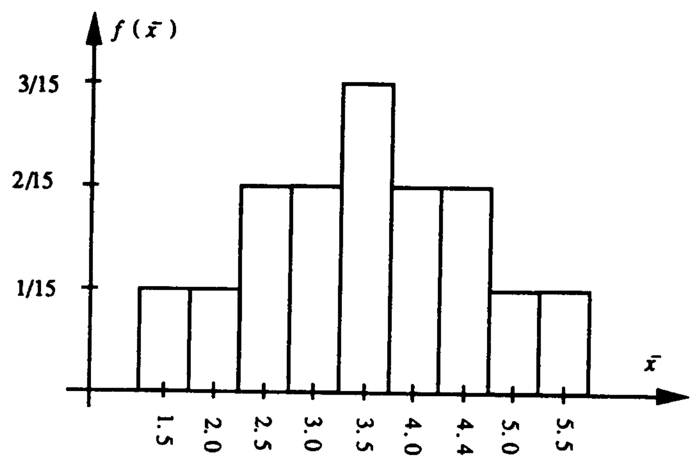
والاحتمال الموافق لـ 2.5 مثلا، هو 1/15 مضروبا بعدد المرات التي تكرر فيها ظهور 2.5 كمتوسط أي:

$$P(\bar{X} = 2.5) = f(2.5) = \frac{1}{15} \times 2 = \frac{2}{15}$$

(انظر الجدول ٤ ـ ٧).

جدول (٤ ـ ٧) التوزيع الاحتمالي لـ $\bar{x}$

$\bar{x}$	1.5	2	2.5	3	3.5	4	4.5	5	5.5
$f(\bar{x})$	1/15	1/15	2/15	2/15	3/15	2/15	2/15	1/15	1/15



شكل (٤ ـ ٥): المدرج الاحتمالي لتوزيع $\bar{x}$ في المثال (٤ ـ ١٣)

جدول (٤ ـ ٦)

العينات الممكنة (نقاط العينة)	الاحتمال الموافق	قيمة $\bar{x}$
1,2	1/15	1.5
1,3	1/15	2
1,4	1/15	2.5
1,5	1/15	3
1,6	1/15	3.5
2,3	1/5	2.5
2,4	1/15	3
2,5	1/15	3.5
2,6	1/15	4
3,4	1/15	3.5
3,5	1/15	4
3,6	1/15	4.5
4,5	1/15	4.5
4,6	1/15	5
5,6	1/15	5.5

لاحظ أن المدرج الاحتمالي لتوزيع $\bar{X}$ متنـاظر تماما في هذا المثال. وهـو مقبب في الوسط ويتناقص تدريجيا على اليمين وعلى اليسار.

(٤ـ٩ـ١) خواص، $\bar{X}$ ، متوسط عينة عشوائية حجمها n مأخوذة من مجتمع حجمه N

لنرمز بـμ (ميو) لمتوسط المجتمع، وبـσ^2 (سيجما مربع) لتباين المجتمع. أي:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \quad , \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2 = \frac{1}{N} \left[\sum_{i=1}^{N} X_i^2 - \frac{\left(\sum_{1}^{N} X_i \right)^2}{N} \right]$$

فيمكن البرهان على الخواص التالية:

١ ـ القيمة المتوقعة لـ $\bar{X}$ تساوي تماما متوسط المجتمع أي:

$$E(\bar{X}) = \mu$$

٢ ـ تباين $\bar{X}$ ولنرمز له بـ $\sigma_{\bar{X}}^2$ معطى بالعلاقة التالية:

$$V(\bar{X}) = \sigma_{\bar{X}}^2 = \frac{N-n}{N-1} \frac{\sigma^2}{n}$$

لاحظ أن تباين $\bar{X}$ الذي يعبر عن تغير قيمة $\bar{X}$ من عينة لأخرى أصغر بكثير من تبـاين المجتمـع. ففي الطـرف الأيمن من عبـارة $V(\bar{X})$ نجــد تبـاين المجتمع σ^2 مقسوما على حجم العينة n، وفوق ذلك أخذنـا جزءا من $\frac{\sigma^2}{n}$ لأن $\frac{N-n}{N-1}$ أصغـر من الواحد. ويكفي أن ننظر إلى المثـال (٤ ـ ١٣) السابق لنجد أن قيم المجتمع تختلف عن بعضها بمقدار الـواحد الصحيح ولكن متوسطـات عينات حجمها 2 مأخوذة من هذا المجتمع لا تختلف عن بعضها إلا بمقدار النصف.

وإذا كــان حجم العينـة n صغيرا جـدا بـالنسبـة إلى حجم المجتمع N تصبح النسبـة $\frac{N-n}{N-1}$ ، وتسمى عامـل التصحيح في مجتمع منته، قـريبـة جـدا من الواحـد ويصبح تباين متوسط العينة مساويا تقـريبا لتباين المجتمع مقسوما على حجم العينة n. وتتضح هنا نقطتان مهمتان:

(١) يمكن التحكم بتباين $\bar{X}$ وجعله صغيرا من خلال زيادة حجم العينة n. أي أن حجم العينة n يشكل صمام أمان تجدر الاستفادة منه للوصول إلى قرارات وتنبؤات إحصائية جيدة. إلا أن مقدرتنا على استخدام صمام الأمان هذا منوطة بالاستعداد لبذل كافة الجهود والنفقات التي يتطلبها أخذ عينة أو يتطلبها القيام بتجربة إحصائية. والمسألة هنا تصبح مسألة اقتصادية إذ نريد، لقاء نفقة معينة، الوصول إلى قرارات أو تنبؤات سليمة، حول خصائص مجتمع، استنادا إلى عينة نأخذها من هذا المجتمع. وتهدف نظرية الاحصاء إلى تقديم طرق كفؤة، تسمح لنا القيام باستقراءات جيدة، من خلال عينات صغيرة نسبيا.

(٢) كلما كان المجتمع المدروس أقل تجانسا (تباينه σ^2 كبير) اضطررنا إلى زيادة حجم العينة حتى نحافظ على حدٍ مرضٍ لسلامة وجودة القرار أو التنبؤ الاحصائي.

مثال (٤ ـ ١٤)

في المثال (٤ ـ ١٣) احسب متوسط المجتمع μ وتباينه σ^2. ثم احسب $E(\bar{X})$ و $V(\bar{X})$ وتحقق من صحة العلاقات المذكورة كخواص للمتوسط.

الحـــل

متوسط المجتمع:

$$\mu = \frac{\sum_{1}^{N} x_i}{N} = \frac{\sum_{1}^{6} x_i}{6} = \frac{1+2+3+4+5+6}{6} = 3.5$$

x_i^2	1	4	9	16	25	36	91
x_i	1	2	3	4	5	6	21

تباين المجتمع:

$$\sigma^2 = \frac{1}{N}\left[\sum_{1}^{N} X_i^2 - \frac{\left(\sum_{1}^{N} X_i\right)^2}{N}\right]$$

$$= \frac{1}{6}\left[91 - \frac{(21)^2}{6}\right] = 2.917$$

$$E(\bar{X}) = \sum_{\bar{x}} \bar{x} \cdot f(\bar{x}) = 1.5 \times \frac{1}{15} + 2 \times \frac{1}{15} + \ldots + 5.5 \times \frac{1}{15} = 3.5 = \mu$$

$$V(\bar{X}) = E(\bar{X})^2 - [E(\bar{X})]^2$$

$$E(\bar{X}^2) = \sum_{\bar{x}} \bar{x}^2 f(\bar{x}) = 1.5^2 \times \frac{1}{15} + 2^2 \times \frac{1}{15} + \ldots + 5.5^2 \times \frac{1}{15} = 13.417$$

$$V(\bar{X}) = 13.417 - 12.25 = 1.167$$

ومن جهة أخرى نجد بتطبيق العلاقة :

$$V(\bar{X}) = \frac{N-n}{N-1}\frac{\sigma^2}{n} = \frac{6-2}{6-1}\frac{2.917}{2} = 1.167$$

وهي النتيجة نفسها التي وجدناها باستخدام التوزيع $f(\bar{x})$ وتعريف التباين.

تمارين (٤ ـ ٥)

١) حظيرة فيها 11 حيوانا منها سبع إناث وأربعة ذكور. اخترنا أربعة حيوانات من الحظيرة عشوائيا. ما احتمالات أن تتضمن:

ا ـ ثلاثة ذكور؟

ب ـ حيوانا واحدا على الأقل من كل جنس؟

٢) نريد اختيار 5 عدائين من بين 10 عدائين ليشكلوا مجموعة أولى A ويشكل الباقون مجموعة B بكم طريقة يمكن القيام بذلك؟ إذا تضمن العداؤون العشرة ثلاثة سعوديين فما احتمالات أن يكون:

ا ـ كل العدائين السعوديين في المجموعة A؟

ب ـ كل العدائين السعوديين في المجموعة نفسها؟
جـ ـ كل مجموعة تتضمن على الأقل عداء سعوديا؟

٣) من صندوق يتضمن 7 قطع معيبة و8 قطع مقبولة، سحبنا عينة عشوائية من خمس قطع، ما احتمال أن تتضمن العينة:

ا ـ قطعة مقبولة واحدة على الأقل؟
ب ـ قطعا من النوعين؟
جـ ـ إذا كانت قيمة كل قطعة مقبولة 15 ريالا وكانت كل قطعة معيبة تسبب خسارة 5 ريالات، فما هي القيمة المتوقعة للعينة؟ وما هو الانحراف المعياري للمتغير الذي يعبر عن قيمة العينة؟
د ـ أعد حل السؤالين (ا) و(ب) بفرض أن السحب مع الإعادة.

٤) من صندوق يتضمن 7 مقاومات ومقاومة كل منها أوم واحد، وثلاث مقاومات، مقاومة كل منها 4 أوم. اخترنا عشوائيا 4 مقاومات ووصلناها على التسلسل لتشكيل وحدة A، وكذلك وصلنا المقاومات الباقية على التسلسل لتشكيل وحدة B. والمقاومة الكلية لوحدة تساوي مجموع مقاوماتها. ما احتمالات:
ا ـ أن تتضمن كل وحدة مقاومة واحدة على الأقل من نوع الـ 4 أوم؟
ب ـ المقاومة الكلية للوحدة B تتجاوز المقاومة الكلية للوحدة A؟
جـ ـ ما المقاومة الكلية المتوقعة لكل من الوحدتين A، B؟

٥) عدد الأسماك N في حوض للأسماك مقدار غير معروف. اصطدنا عشرين سمكة من الحوض ووضعنا على كل منها علامة مميزة ثم أعدناها إلى الحوض. ثم عدنا فاصطدنا 25 سمكة ووجدنا أنها تتضمن 3 سمكات معلّمة. عبر بدلالة N عن احتمال هذه الحادثة، ولنرمز لهذا الاحتمال بـ P_N. ويعتبر تقديرا جيدا لـ N، تلك القيمة التي تجعل P_N أكبر ما يمكن. وبملاحظة أن:

$$\frac{P_N}{P_{N-1}} = \frac{N^2 - 45N + 500}{N^2 - 42N}$$

نجد أن $P_{N-1} < P_N$ إذا، وفقط إذا، كـان $N > 500/3$. ما هو تقـديرك لعدد الأسماك في الحوض؟

٦) خذ عينة حجمها 4 من المجتمع المذكور في المثال (٤ ــ ١٣) واكتب توزيع $\bar{X}$ ، ثم احسب $E(\bar{X})$و $V(\bar{X})$. ما أثر زيادة حجم العينة على تباين $\bar{X}$ وعلى شكل المدرج الاحتمالي لتوزيع $\bar{X}$ ؟

٧) من بين 15 متقدما لوظيفة ما، تسعة منهـم يحملون درجة جامعية. اختير متقدمان عشوائيا لإجراء مقابلة. أوجد احتمال أن:
ا ـ أحدهم يحمل درجة جامعية والآخر لا يحمل درجة جامعية.
ب ـ ليس بينهما من يحمل درجة جامعية.
جـ ـ كلاهما يحمل درجة جامعية.

٨) لدى سكـرتير 9 رسائل وعليـه أن يرسل ثـلاثا منهـا محددة بالبريد المسجـل والباقي بالبريد العادي. اختلط عليه الأمـر فاختار عشوائيا 3 رسائل ووضع عليهـا طوابع البريد المسجل. ما هو احتمال:
ا ـ أن اختياره لم يكن صحيحا تماما،
ب ـ أن اختياره كان صحيحا.

٩) بالاشـارة إلى التمريـن ١٣ من مجمــوعة التمـارين (١ ـ ١) ، لنعتبر أن القيـاسات الخمسمائة تشكل مجتمعـا من قيـاســات معـدل الكـولسترول في الـدم. ولنسجل الأعـداد الخمسمائة (إذا أمكن)، أو مـا يشير إليهـا، على قطع صغيرة من الـورق، ولنختر عشوائيا عينـة حجمها عشرة باختيارنا عشـوائيا لعشرة أوراق (مع الاعادة). لنكرر تجربـة سحب العينة هذه مـائة مرة ولنحسب المتوسطـات المائة لهذه العينات ونرسم لها مدرج تكـرار نسبي مستخدمين حدود الفئات نفسهـا. قارن الآن مدرج التكـرار النسبي الحاصل مع المدرج الخاص بـالمجتمع المطلـوب في الجزء جــ من التمرين ١٣ من مجموعة التمـارين (١ ـ ١). هل تجده أقرب إلى مدرج المجتمع من

المدرجات المطلوبة في الجزء ب من التمرين ١٣ . احسب متوسط البيان الاحصائي من مائة متوسط وتباينه وقارنهما مع متوسط المجتمع وتباينه. ماذا تستنتج؟ (من المفضل أن يتعاون الفصل بكامله في حل هذا التمرين) .

الفصل الخامس

التوزيع الطبيعي

(٥ ـ ١) مقدمة

رأينا في الفقرة (٣ ـ ٥) أن المتغيرات العشوائية المستمرة تولد فضاء عينة مستمرا، بمعنى أن نقاطه تكون متراصة بعضها إلى بعض كنقاط محور موجه. وبالتالي فإنها، بالإضافة إلى كونها لا نهائية في عددها، غير قابلة للعد. وكأمثلة تقليدية على متغيرات عشوائية مستمرة، نذكر أطوال البشر وأوزانهم، وأخطاء القياس في تجربة مخبرية، وعمر مصباح كهربائي، إلخ. كما رأينا في تلك الفقرة أنه للحصول على نموذج احتمالي لمتغير عشوائي مستمر، X، نبدأ باختيار منحن مستمر يمثل ما سميناه بدالة الكثافة الاحتمالية، وأن مثل هذه الدالة، ولنرمز لها بـ $f(x)$، يجب أن تحقق شرطين:

١ ـ $f(x) \geq 0$، مهما يكن x،
٢ ـ المساحة تحت $f(x)$ تساوي الواحد تماما.

وعندئذ يكون احتمال أي حادثة عددية مثل $a < X < b$، حيث a، b عددان محددان، هو المساحة تحت منحنى الكثافة فوق الفترة (a, b) من محور السينات. وكنتيجة لذلك نجد أن احتمال أن يفترض المتغير X قيمة معينة، a مثلا، أي $P(X = a)$، هو المساحة تحت المنحنى فوق النقطة a، وهي صفر. وهكذا فإن مثل

هذا الحل لمشكلة إيجاد نموذج احتمالي لفضاء عينة مستمر يحتم علينا القول إن احتمال أن يكون لمتغير عشوائي مستمر قيمة معينة هو احتمال يساوي الصفر. وهذا تعبير واقعي عن استحالة توصل الإنسان إلى أجهزة قياس دقيقة بصورة مطلقة.

وبينما تتخذ منحنيات الكثافة أشكالا مختلفة نلاحظ أن عددا كبيرا من المتغيرات العشوائية التي نواجهها في حياتنا العامة لها منحني كثافة، أو منحني تكرار، له تقريبا شكل الجرس، أو، كما نعبر عن ذلك إحصائيا، له بصورة تقريبية شكل منحني التكرار الطبيعي، أو شكل التوزيع الطبيعي.

وبصورة عامة لنفرض أننا لاحظنا ، في مجتمع القياسات لظاهرة معينة، ميلا واضحا إلى التناظر والاعتدال، بمعنى أن القياسات المتطرفة التي تمثل فرط زيادة أو فرط نقصان، هي قياسات نادرة. ويزداد تكرار ظهور القياس في ذلك المجتمع كلما اقتربت قيمة القياس من المتوسط. فالقيمة المتوسطة في المجتمع والقيم المجاورة لها هي القياسات الأكثر تواترا، بينما تكون القياسات البعيدة عن المتوسط زيادة أو نقصانا نادرة الظهور. وبعبارة أخرى ، لنفرض أن الوسطية والاعتدال هي السائدة في مجتمع القياسات لظاهرة معينة، فعندئذ نقول إن النموذج الاحتمالي المناسب لهذه الظاهرة هو نموذج «التوزيع الطبيعي». وقد برزت تسمية «الطبيعي» في القرن الثامن عشر في سياق نظرية «أخطاء القياسات» عندما وجد أنه في تجربة يسير كل شيء فيها سيرا طبيعيا (normally) ، ستكون أخطاء القياسات خاضعة للتوزيع الاحتمالي الذي يتخذ منحنى الكثافة فيه شكل الجرس (أو شكل منحنى جاوس). وتجدر هنا ملاحظة أنه عندما تتوافر كفاءة المجرب ومقدرته على إجراء القياسات بصورة سليمة، وتتوافر إلى جانب ذلك سلامة الأجهزة المستخدمة، وسلامة الظروف التي تتم تحتها التجربة، فإن الأخطاء ستتذبذب بصورة قريبة من التناظر بين أخطاء بالزيادة وأخطاء بالنقصان، وستكون الأخطاء الفاحشة بالزيادة أو بالنقصان نادرة، بينما تتمركز معظم نتائج القياسات حول القيمة الحقيقية، التي تشكل المتوسط، وقريبا منها. وينبغي ألا تملي التسمية أي شكل من أشكال خصوصية هذا التوزيع لعلوم الطبيعة، فهو يلعب، في

الـواقع، دورا أعم من ذلك بكثير وأوسع، وهو بين التـوزيعات الاحتمالية، بمختلف أنـواعهـا ومسميـاتها، علم بـارز، إليـه تستند، بصـورة رئيسـة، العـديـد من الطـرق الإحصـائية، وبـدونه تضيق الحلبـة الواسعـة لتطبيقات الإحصـاء في الحياة المعـاصرة. وسنجـد فيما يسمى «نظـريـة النهـايـة المركـزيـة» أن مجمـوع عـدد كبير من المركَبـات العشـوائية، هـو دائما متغير عشـوائي ينزع، تحت شـروط عـامة جدا، إلى الخضـوع للتوزيـع الطبيعي، وذلـك بصرف النظر عن طبائع تلـك المركبات العشوائيـة التي تمثل كل منها متغيرا عشـوائيا لـه توزيعه الاحتمالي الخاص. وقـد رأينا في الفصـل الثـاني أن المعـايير الإحصائية المهمة يعبر عنهـا بدلالـة مجمـوع متغيرات، فمثلا، $n\bar{x} = \sum_{1}^{n} x_i$ و $(n-1)\,S^2 = \sum_{1}^{n} (x_i - \bar{x})^2$، كما رأينا في الفصل الرابـع أن عدد النجاحات، X، في تجربة ثنائية ما هـو إلا مجموع عينة حجمها n مأخوذة من مجتمع بيرنوللي. وهذا يشير بوضوح إلى الأهمية الخاصة لهذا التوزيع في مباحث الإحصاء.

(٥ ـ ٢) دالة الكثافة الاحتمالية للتوزيع الطبيعي

تعرف دالة الكثافة الإحتمالية للتوزيع الطبيعي كما يلي:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad \begin{array}{l} -\infty < x + \infty \\ -\infty < \mu < +\infty \\ 0 < \sigma < +\infty \end{array}$$

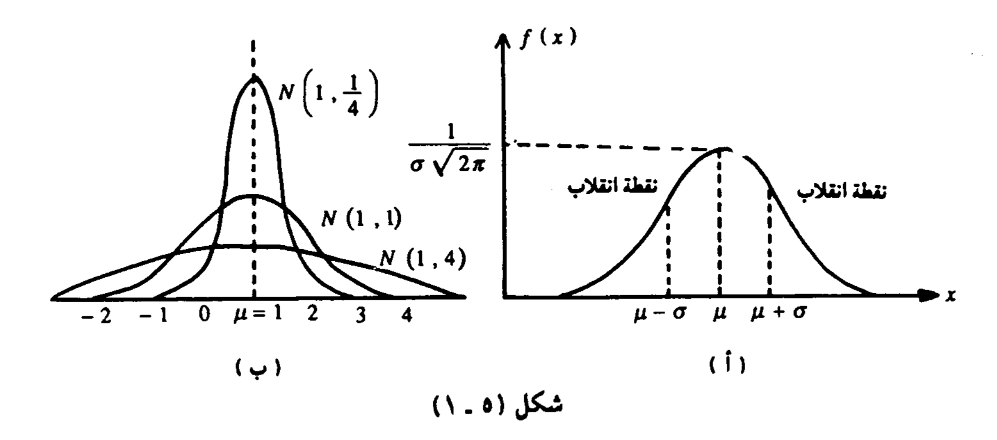
وهي دالة منحن له شكل الجرس (انظر الشكل ٥ ـ ١ (أ)) حيث:

π عدد ثابت يساوي تقريبا 3.1416 ،

e عدد ثابت يساوي تقريبا 2.7183 ،

μ عدد ثابت يمكن أن يكون أي عدد حقيقي ،

σ عدد ثابت يمكن أن يكون أي عدد حقيقي موجب.



شكل (٥ ـ ١)

والدالة أعلاه لا تحدد منحنيا واحدا بعينه و إنما تحدد الشكل العام لعائلة من المنحنيات . إذ كلما حددنا لـ μ قيمة ولـ σ قيمة نحصل على منحن محدد تماما . ولذلك يسمى كل من الثابتين μ ، σ مَعلمة .

ويمكن البرهان على أن المعلمة μ تمثل متوسط التوزيع الاحتمالي ، أي $E(X) = \mu$ ، وأن المعلمة σ تمثل الانحراف المعياري للتوزيع الاحتمالي ، أي $V(X) = \sigma^2$. وللمنحنيات الطبيعية المختلفة متوسطات مختلفة ، وانحرافات معيارية مختلفة ، إلا أن المتوسط μ والانحراف المعياري σ لمنحن طبيعي معين محددان تمام التحديد وثابتان . وهكذا نجد أن تحديد قيمة لـ μ وقيمة لـ σ يحدد تماما منحنيا ، وعلى العكس كل منحن من عائلة المنحنيات الطبيعية (منحنيات جاوس أو المنحنيات على شكل جرس) تحدد تماما قيمة لـ μ وقيمة لـ σ . وهذا يلقي بعض الضوء على سبب تسمية μ و σ مَعلمات . ويُبرهَن في الحساب التكاملي أن المساحة تحت المنحنى

$$e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

يساوي $\sigma\sqrt{2\pi}$ تماما . وبالتالي فإن المساحة تحت المنحنى الطبيعي $f(x)$ ، كما عرفناه أعلاه تساوي الواحد تماما .

ونلاحظ أن المنحني متناظر حول المستقيم $X = \mu$ الموازي للمحور الرأسي . لأن الدالة f تأخذ القيمة نفسها في نقطتين متناظرتين بالنسبة إلى النقطة $x = \mu$ ، فلو حسبنا ، مثلا ، $f(\mu + a)$ و $f(\mu - a)$ لوجدنا :

$$f(\mu + a) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left[\frac{\mu + a - \mu}{\sigma}\right]^2} = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\frac{a^2}{\sigma^2}}$$

$$f(\mu - a) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left[\frac{\mu - a - \mu}{\sigma}\right]^2} = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\frac{a^2}{\sigma^2}}$$

فالنقطة $x = \mu$ على المحور الأفقي هي النقطة التي يتمركز عندها التوزيع $(E(X) = \mu)$ ، وينتشر على جانبيها بصورة متناظرة .

ومن دراستك السابقة للدالة الرأسية e^{-t} ، مثلا ، تذكر أنه إذا كان الأس t موجبا دوما ، كما هو الحال في الدالة $f(x)$ هنا حيث $t = \frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2$ ، فإن أكبر قيمة لـ e^{-t} تساوي الواحد ، وهي القيمة الموافقة لـ $t = 0$ ، (في مثالنا $x = \mu$) . وتتناقص قيمة e^{-t} مع تزايد t وتنتهي إلى الصفر (أي تتقارب إلى المحور الأفقي) عندما تزداد t إلى اللانهاية .

وهكذا فإن دالة الكثافة $f(x)$ تبلغ نهايتها العظمى عند $x = \mu$ وتكون قيمتها عندئذ :

$$f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{\mu - \mu}{\sigma}\right)^2} = \frac{1}{\sigma\sqrt{2\pi}}\ e^0 = \frac{1}{\sigma\sqrt{2\pi}}$$

وللمنحنى $f(x)$ نقطتا إنقلاب عند $x = \mu + \sigma$ و $x = \mu - \sigma$ (أنظر الشكل ٥ ـ ١ (أ)) . لنتصور في الشكل ٥ ـ ١ (أ) أن المنحنى عبارة عن سلك رفيع وشديد المرونة . فإذا ضغطنا على القمة سينتشر السلك انتشارا أوسع على جانبي μ ، أي يأخذ شكلا أكثر انبساطا باعتبار أن المساحة تحت السلك يجب أن تبقى دائما ثابتة ومساوية للواحد .

وإذا رفعنا القمة إلى أعلى فسيقل انبساط المنحنى ويتضاءل انتشاره على جانبي μ .

ونرى في الشكل ٥ ـ ١ (ب) تمثيلا يوضح الفكرة . وقد استخدم الرمز $N(\mu, \sigma^2)$ للدلالة على توزيع طبيعي بمتوسط μ وتباين σ^2 ، وهكذا يعني $N\left(1, \frac{1}{4}\right)$ توزيعا طبيعيا متوسطه $\mu = 1$ وتباينه $\sigma^2 = \frac{1}{4}$ ، وللمنحنيات الثلاثة في الشكل (٥ ـ ١) ب المتوسط نفسه وهو 1 . وعندما ارتفعت قمة المنحنى $N(1, 1)$ تضاءل انتشاره على جانبي المتوسط $\mu = 1$ وبالتالي قل σ^2 من 1 إلى $\frac{1}{4}$ ، وعلى العكس عندما انخفضت قمة المنحنى ، اتسع

انتشـاره على جـانبي μ وازداد تباينه من 1 إلى 4 . ولو عــدنا إلى قيمـة $f(x)$ العظمـى وهي $f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$ لتبين لنـا أن القيم الصغـيرة لـ σ تعني قمـة مرتفعة، أي توزيعـا أقل انتشارا حـول متوسطـه، وأن القيم الكبيرة لـ σ تعني قمـة منخفضة، أي توزيعا أكثر انتشارا على جـانبي المتوسط . ولما كان التباين، كما نعلم من الفصلين الثاني والرابع، مقياسا لمدى انتشار التوزيع على جانبي المتوسط، فإن هذه الملاحظة توضح أن σ^2 يمثل تبـاين التـوزيع الأمر الـذي ذكـرناه منـذ قليل كنتيجـة يمكن إثبـاتها رياضيـا باستخدام الحساب التكاملي وبطرق تعتبر فوق مستوى هذا الكتاب .

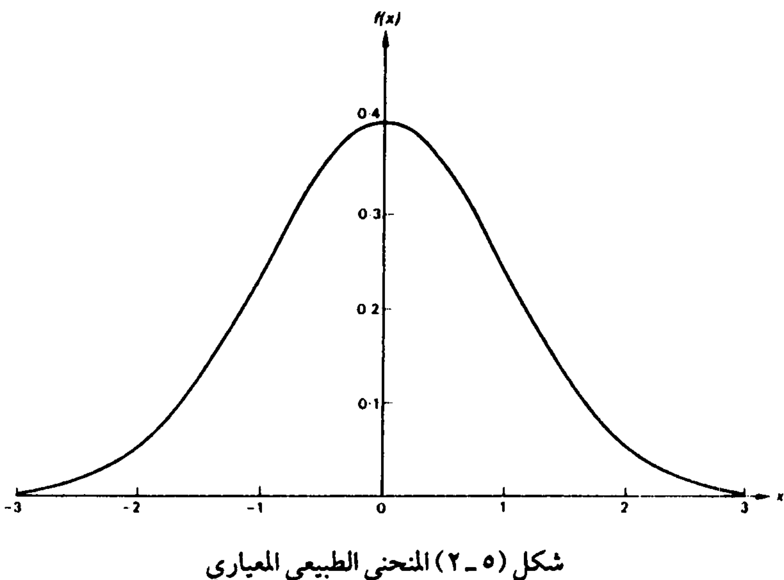
ومن بين أسرة المنحنيات الطبيعية :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad \begin{array}{l} -\infty < x + \infty \\ -\infty < \mu < +\infty \\ 0 < \sigma < +\infty \end{array}$$

سنختار منحنيـا خاصا هـو ذلك المنحنى الذي يكون متـوسطه $\mu = 0$ وانحرافه المعياري $\sigma = 1$. وتمييـزا لهذا المنحنى، الذي سيلعب دورا هـامـا في تطبيقات التـوزيع الطبيعي سنطلق عليــه اسم المنحنى الطبيعي المعيـاري . وإذا استخــدمنـا الحرف Z للمتغير الطبيعي المعياري فستصبح معادلة المنحنى أعلاه بعد وضع $\mu = 0$ ، $\sigma = 1$ على الشكل

$$f(Z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}Z^2} \quad ; \quad -\infty < Z < +\infty \; .$$

ونجد في الشكل (٥ ـ ٢) الرسم البياني لهذا المنحنى . وتجدر ملاحظة أنه متناظر بالنسبـة إلى المحـور الرأسـي . ومـا دامت المساحة تحت المنحنى بكامله من $Z = -\infty$ إلى $Z = +\infty$ هـي الواحد تمامـا فالمساحة على اليمين من $Z = 0$ تساوي المساحة على اليسار من $Z = 0$ وكل منهما تساوي النصف .



شكل (٥ ـ ٢) المنحنى الطبيعي المعياري

تمارين (٥ ـ ١)

١) اكتب دالة الكثافة للتوزيع الطبيعي مفترضا القيم التالية للمتوسط والتباين :

أ ـ المتوسط يساوي 3 ، والتباين يساوي 4 .

ب ـ المتوسط يساوي 0 ، والتباين يساوي 5 .

جـ ـ المتوسط يساوي 2− ، والتباين يساوي 1 .

د ـ المتوسط يساوي 6− ، والتباين يساوي 10 .

حدد في كل حالة أين تقع قمة المنحنى وحاول أن تخطط رسما تقريبيا له .

٢) متغير عشوائي X دالة كثافته الاحتمالية :

$$f(x) = \frac{1}{3\sqrt{2\pi}} \; e^{-\frac{1}{18}x^2} \quad ; \quad -\infty < x < +\infty \; .$$

ما متوسطه و انحرافه المعياري؟

٣) متغير عشوائي X دالة كثافته الاحتمالية :

$$f(x) = c\, e^{-\frac{(x-4)^2}{6}} \quad ; \quad -\infty < x < +\infty .$$

ما قيمة c ؟

. (٥ ـ ٣) المساحات تحت منحنى الكثافة الطبيعي

ذكرنا أن معادلة منحنى الكثافة الطبيعي، كما وردت في مستهل الفقرة السابقة، لا تمثل منحنيا واحدا، بل عائلة من المنحنيات لا حصر ولا عد لأعضائها. ووضع جدول للمساحات خاص بكل منها أمر غير ممكن. وسنجد الآن أنه يمكن وضع جدول واحد كاف لحساب المساحات تحت أي منحنى كثافة طبيعي. وأسهل طريقة لتحقيق ذلك هي أن نحسب المساحات الواقعة ضمن عدد محدد من الانحرافات المعيارية على جانبي المتوسط. وبما أن المنحنى متناظر يمكن التبسيط بإقامة جدول للمساحات تحت المنحنى بين μ والنقاط x الواقعة على اليمين من μ . وإذا فرضنا نقطة x أكبر من μ فإن المسافة بين x و μ هي $x-\mu$ ، وإذاعبرنا عنها بدلالة الانحراف المعياري σ ، ولنفرض أنها تساوي Z مرة الانحراف المعياري σ ، فيمكننا أن نكتب $x-\mu = Z\sigma$ ، وإذا قسنا المسافات على محور الفواصل بوحدة قياس تساوي σ (وعندها يكون $\sigma = 1$ حكما) فإن قيمة المسافة $x-\mu$ مقيسة بالوحدة الجديدة تصبح Z أي تساوي $\frac{x-\mu}{\sigma}$. وهكذا نكتب المتغير الجديد Z بدلالة المتغير X على الشكل :

$$Z = \frac{X-\mu}{\sigma}$$

ونلاحظ أنه يوافق كل قيمة لـX قيمة واحدة لـZ والعكس بالعكس. وأن Z ليس إلا القيمة المعيارية لـ X. وفي الواقع، لو حسبنا $E(Z)$ و $V(Z)$ لوجدنا :

$$E(Z) = E\left[\frac{1}{\sigma}(X-\mu)\right] = \frac{1}{\sigma}\, E(X-\mu) = \frac{1}{\sigma}\,[\,E(X)-\mu\,] = 0$$

$$V(Z) = V\left(\frac{1}{\sigma} \times -\frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2}\quad V(X) = \frac{\sigma^2}{\sigma^2} = 1 .$$

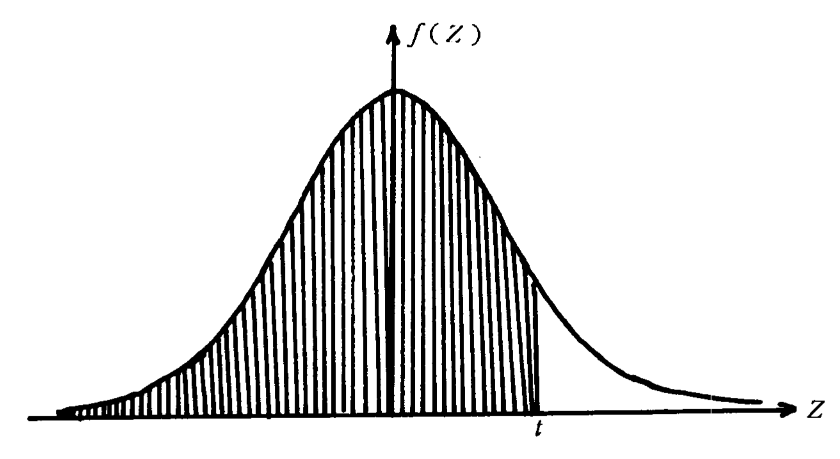
أي أن للمتغير Z متوسطا يساوي الصفر وانحرافا معياريا يساوي الواحد، ويمكن البرهان على أن التوزيع الاحتمالي لـZ هو التوزيع الطبيعي. وبذلك يكون

منحنى الكثــافة الموافق لــZ عضــوا في أسرة المنحنيــات الطبيعيــة ، وبــالــذات ذلك العضو المقابل لـ$\mu = 0$ و$\sigma = 1$. وهو بالضبط منحنى الكثافـة المذكور في ختام الفقرة (٥ـ٢):

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} Z^2} \quad ; \quad -\infty < Z < +\infty .$$

وربما أصبح واضحا الآن سبب تسمية هذا المنحنى بالمنحنى الطبيعي المعياري.

ويقـدم جدول التـوزيع الطبيعي في الملحق، المسـاحات تحت هـذا المنحنى إلى اليسار من نقطة معينة$Z = t$. ونقصد المساحة المظللة في الشكل (٥ ـ ٣).



شكل (٥ ـ ٣) دالة التوزيع المتجمع للمتغير الطبيعي المعياري Z

وبالعودة إلى جدول التـوزيع الطبيعي المشار إليه في الملحق، نلاحظ أن قيم Z في الجدول تبدأ مـن الصفر بفاصـل قدره 0.01 بين كل قيمـة والقيمة التي تليهـا. وهكذا تكون كل قيمـة لـZ معطاة بـرقمين عشريين. ويتضمن العمـود الأول قيما لـZ بفاصل يساوي 0.1 من قيمة إلى القيمة التي تليها. وتشكل هذه القيم عناوين لسطور الجدول، إذ نبدأ بالسطر 0 يليه السطر 0.1 ، فالسطر 0.2 ، وهكذا حتى نصل إلى السطر 3.4 . أما المنزلـة العشرية الثانيـة من قيمة Z فهي معطــاة في السطــر الأفقي الأول من الجدول ، وتشكل عناوين لأعمـدة الجدول، بدءا من العمود الثاني حتى العمـود الأخير، وهكذا

نجد العمود 0.00 يليه العمود 0.01 ، يليه العمود 0.02 ، وهكذا حتى نصل إلى العمود 0.09 وهو العمود الأخير. وكل عدد في صلب الجدول، وهو ملتقى سطر مع عمود، يمثل المساحة تحت منحنى الكثافة المعياري وإلى اليسار من قيمة Z التي يحددها عنوان السطر حتى الرقم العشري الأول ويستكمل عنوان العمود رقمها العشري الثاني. وهكذا فإن العدد 0.8212 الواقع في ملتقى السطر 0.9 مع العمود 0.02 ، يمثل المساحة تحت منحنى الكثافة الواقعة إلى اليسار من $Z = 0.92$ ، أي المساحة تحت المنحنى وفوق الفترة الممتدة بين $\infty -$ والنقطة 0.92 من المحور Z . وعلى العكس، إذا أردنا المساحة الواقعة إلى اليسار من $Z = 1.96$ ، مثلاً، ندخل الجدول وفق السطر 1.9 والعمود 0.06 فنجد عند ملتقاهما العدد 0.9750 وهو المساحة المطلوبة. وإذا كانت قيمة Z معطاة بأكثر من رقمين عشريين فإننا نحصرها بين قيمتين مذكورتين في الجدول ثم نقوم بعملية تناسب طردي، (عملية استيفاء).

والأسئلة الوجيهة التي تطرح نفسها هنا هي:

١ ــ إذ يقتصر الجدول على القيم الموجبة لـ Z ، ما العمل لو كانت القيمة المعطاة لـ Z سالبة؟

٢ ـ ما العمل لو كان المطلوب هو المساحة إلى اليمين من قيمة لـ Z سالبة أو موجبة؟

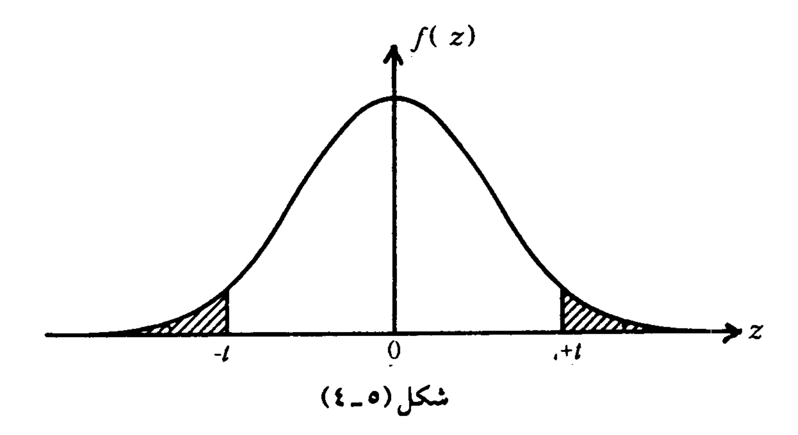
٣ ـ ما العمل لو كان المطلوب هو المساحة بين أي قيمتين لـ Z ؟

وللإجابة عن هذه التساؤلات نعود إلى التعريف في (٣ ـ ٦ ـ ٢) لدالة التوزيع المتجمع، ونكتب: المساحة تحت المنحنى الطبيعي المعياري الواقعة إلى اليسار من النقطة t = $P(Z \leq t) = F(t)$

وتتمتع هذه الدالة $F(t)$ بالخاصة المهمة التالية:

$$F(-t) = 1 - F(t)$$

وهي نتيجة مباشرة لتناظر المنحنى الطبيعي المعياري بالنسبة إلى المحور الرأسي. إذ لو نظرنا إلى الشكل (٥ ـ ٤) لوجدنا أن $F(-t)$ يساوي المنطقة I المظللة على اليسار من



شكل (٥ ـ ٤)

$Z = - t$. وأن $F(t)$ هي مجموع المنطقـة المظللة في أقصى اليسار والمنطقـة غير المظللة في الوسط . و $1 - F(t)$ يساوي بوضـوح المنطقة المظللة في أقصى اليمين، وبما أن المنطقتين المظللتين متساويتان بحكم التناظر فإن $F(- t) = 1 - F(t)$. ولإيجاد $F(- t)$ يكفي إذن حساب $F(t)$ من الجدول الموصوف أعلاه، حيث t موجبة، ثم نطرح القيمة الناتجة من 1 وهذا يجيب عن السؤال الأول.

ومن خاصة الحادثتين المتتامتين، $P(A) = 1 - p(\bar{A})$ ، نجد مباشرة أن:

$$P(Z > t) = 1 - P(Z \leq t) = 1 - F(t)$$

وهذا يجيب عن السؤال الثاني.

وللإجابة عن السؤال الثالث، لنفرض أن المطلوب هو حساب

$$P(a < Z \leq b)$$

فمن الواضح أنـه يمكن التعبير عن الحادثة $(Z \leq b)$ كإتحاد حـادثتين منفصلتين على الشكل

$$(Z \leq b) = (Z \leq a) \cup (a < Z \leq b)$$

ومنه:

$$P(Z \leq b) = P(Z \leq a) + P(a \leq Z \leq b)$$

أي:

$$F(b) = F(a) + P(a < Z \leq b)$$

أو

$$P(a < Z \leq b) = F(b) - F(a)$$

وبما أن الاحتمال الموافق لنقطة في التوزيعات المستمرة يساوي الصفر فيمكن كتابة

$$P(a \leq Z < b) = P(a < Z \leq b) = P(a \leq Z \leq b) = P(a < Z < b)$$

مثال (٥ ـ ١)

احسب $P(Z \leq 1.35)$ ، $P(Z > 0.5)$ ، $P(Z < -1.79)$ ، $P(Z \geq -0.68)$ ، $P(1 < Z < 3.27)$ ، $P(-0.1 < Z < 2.5)$ ، $P(-1.85 < Z < -0.16)$.

الحل

$$P(Z \leq 1.35) = F(1.35) = 0.09115$$

(ندخل الجدول وفق السطر 1.3 والعمود (0.05) .

$$P(Z > 0.5) = 1 - F(0.5) = 1 - 0.6915 = 0.3085$$

(ندخل الجدول وفق السطر 0.5 والعمود 0.00) .

$$P(Z < -1.79) = F(-1.79) = 1 - F(1.79) = 1 - 0.9633 = 0.0367$$

(ندخل الجدول وفق السطر 1.7 والعمود 0.09) .

$$\mathrm{P}(Z \geq -0.68) = 1 - \mathrm{P}(Z < -0.68) = 1 - F(-0.68)$$

$$= 1 - [1 - F(0.68)] = F(0.68) = 0.7517$$

(ندخل الجدول وفق السطر 0.6 والعمود 0.08) .

$$P(1 < Z < 3.27) = F(3.27) - F(1)$$

$$= 0.9995 - 0.8413 = 0.1582$$

$$P(-0.1 < Z < 2.5) = F(2.5) - F(-0.1)$$

$$= F(2.5) - [1 - F(0.1)]$$

$$= F(2.5) + F(0.1) - 1$$

$$= 0.9938 + 0.5398 - 1 = 1.5336 - 1 = 0.5336$$

$$P(-1.85 < Z < -0.16) = F(-0.16) - F(-1.85)$$
$$= [1 - F(0.16)] + [1 - F(1.85)]$$
$$= 2 - F(0.16) - F(1.85)$$
$$= 2 - 0.5636 - 0.9678 = 2 - 1.5314 = 0.4686$$

لاحظ أننا نعود إلى الجدول عندما يكون المطلوب $F(t)$ حيث t عدد موجب .

مثال (٥ ـ ٢)

احسب c بحيث يكون

$P(Z \leq c) = 0.8264$ ، $P(Z < c) = 0.2981$ ، $P(Z > c) = 0.9292$ ،
$P(-c < Z < c) = 0.9500$ $P(-c < Z < c) = 0.90$.

الحل

$$P(Z \leq c) = F(c) = 0.8264$$

والعدد c هو قيمة Z في جدول التوزيع الطبيعي المعياري التي يقع إلى اليسار منها مساحة تساوي 0.8264 . ونبحث في صلب الجدول عن هذه القيمة لنجدها بالذات وعندئذ نحدد قيمة Z المطلوبة من السطر والعمود الموافقين ، أو نحصرها بين عددين في الجدول ثم نستنتج قيمة Z المطلوبة بعملية تناسب طردي (استيفاء) . وفي حالتنا هنا نجد أن 0.8264 واقع في السطر 0.9 والعمود 0.04 وتكون القيمة c المطلوبة 0.94 .

$$P(Z < c) = 0.2981 \Leftrightarrow F(c) = 0.2981$$

وإذا كانت قيمة $F(c)$ أصغر من 0.5 فمن الواضح أن c ستكون سالبة . ولكن الجدول لا يحوي القيم السالبة لـ Z . وفي مثل هذه الحالة نأخذ :

$$F(-c) = 1 - F(c) = 1 - 0.2981 = 0.7019$$

ونبحث في صلب الجدول عن 0.7019 فنجده في السطر 0.5 والعمود 0.03 وتكون $-c = 0.53$ أو $c = -0.53$.

$$P(Z > c) = 0.9292 \Leftrightarrow 1 - F(c) = 0.9292$$

أي

$$F(-c) = 0.9292 \Leftrightarrow -c = 1.47 \Leftrightarrow c = -1.47$$
$$P(-c < Z < c) = 0.95 \Leftrightarrow F(c) - F(-c) = 0.95$$

ومنه

$$F(c) - [1 - F(c)] = 0.95$$

$$2F(c) = 1.95,\ F(c) = 0.975,\ c = 1.96.$$

$$P(-c < Z < c) = 0.90 \Leftrightarrow 2F(c) = 1.90$$

أي

$$F(c) = 0.95$$

ولدينا من الجدول

$Z = 1.64$ تقابل 0.9495

$Z = 1.65$ تقابل 0.9505

ومنه

تزايد المساحة	تزايد Z
0.001	0.01
0.0005	?

$$\text{التزايد المطلوب في Z} = \frac{0.0005 \times 0.01}{0.001} = 0.005$$

وتكون قيمة Z المطلوبة هي

$$1.64 + 0.005 = 1.645$$

وسنصطلح على كتابة Z_α لتعني قيمة المتغير الطبيعي المعياري التي يقع إلى اليمين منها مساحة تساوي α. أي أن $F(Z_\alpha) = 1 - \alpha$. وبهذا المعنى يكون :

$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = F(Z_{\alpha/2}) - F(-Z_{\alpha/2})$$

$$= 2F(Z_{\alpha/2}) - 1 = 2\left(1 - \frac{\alpha}{2}\right) - 1 = 1 - \alpha$$

وعلى سبيل المثال :

$Z_{.025}$ هي قيمة Z التي تحصر إلى اليمين منها مساحة تساوي 0.025. ويكون

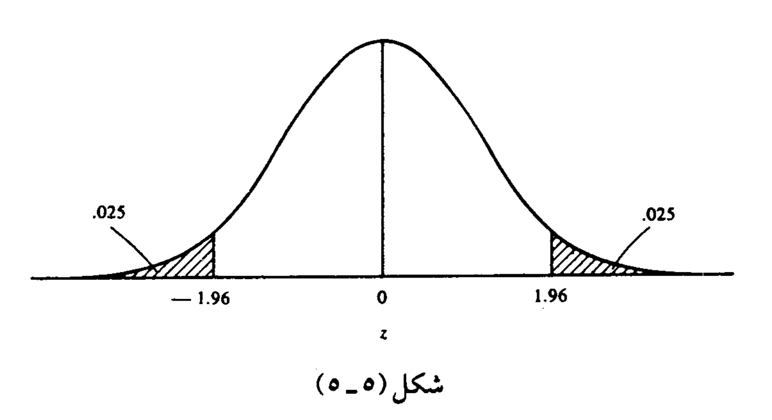
$$P(-Z_{0.025} < Z < Z_{0.025}) = 0.95$$

وقد رأينا في المثال السابق أن



$$Z_{0.025} = 1.96$$

(أنظر الشكل ٥ ـ ٥).



شكل (٥ ـ ٥)

لقد تعلمنا حتى الآن كيف نحسب احتمالات حوادث معبرا عنها بدلالة المتغير المعياري Z ، وذلك بالاستفادة من جدول التوزيع الطبيعي المعياري. ولكن كيف نستفيد من هذا الجدول نفسه لحساب احتمالات حوادث معبر عنها بدلالة متغير طبيعي غير معياري، X ، مثلا؟ بالطبع لا يمكننا حساب مثل هذه الاحتمالات إلا إذا حددنا منحنى الكثافة للمتغير X تحديدا تاما. أي علمنا متوسطه μ وانحرافه المعياري σ . وعند معرفة قيمة μ وقيمة σ يصبح الأمر في غاية السهولة، إذ نقوم بمعايرة X ، أي نكتب:

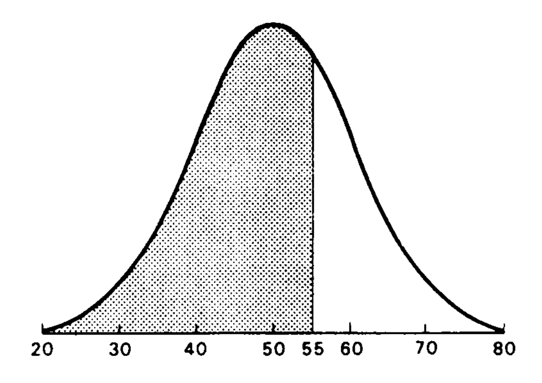
$$Z = \frac{X - \mu}{\sigma}$$

ونحول العبارة الاحتمالية بدلالة X إلى عبارة احتمالية مكافئة بدلالة Z ، ثم نعود إلى جدول التوزيع الطبيعي المعياري، الذي تدرَّبنا لتونا على كيفية استخدامه، لحساب المطلوب وفيما يلي توضيح عملي للفكرة.

يقدم اختصاصي في علم النفس نصائح حول أفضل المهن أو الوظائف المناسبة لفتى. ولهذه الغاية يقدم للفتى عددا من الاختبارات. أحدها ، مثلا، اختبار يهدف إلى قياس مهارات التحدث أو المهارات الشفهية. لنفرض أن درجة الفتى في هذا

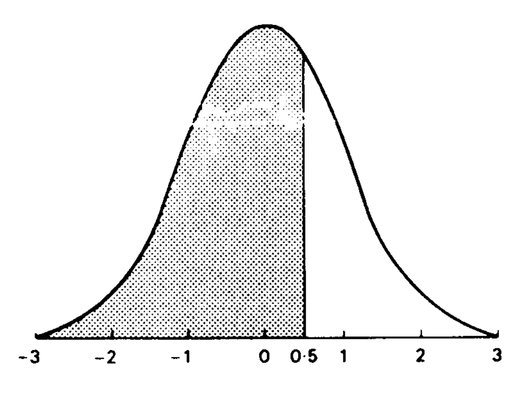
الاختبار كـانت 55 . فهـذا الـرقم لـذاتـه ليس لـه أي مدلـول بـالنسبـة إلينـا . إلا أن الاختصـاصي النفسي يعلم توزيع درجـات هـذا الاختبار بـالنسبـة للرجـال في المجتمع بصـورة عــامة . فمثل هـذه الاختبـارات قـد استخـدمت في الماضي على نطـاق واسع وقدمت لعيــنة تمثيليـة كبيرة من الرجال والنســاء . وبالنسبة إلى الرجـال تتوزع درجات هذا الاختبــار، بصورة تقـريبية، وفـق التوزيع $N(50, 10^2)$. (في الـواقع يتعمـد مصممـو هـذه الاختبارات وضعهـا بحيث تتوزع الدرجـات الناتجة عنهـا طبيعيا، على وجه التقريب) . وتوزيع هذه الدرجـات مع درجة الفتى مبينة في الشكل (٥ ـ ٦) . وما يهم الاختصـاصي النفسي حقا هــو كيف يمكـن مقـارنة هـذا الرجـل مع بقيـة الرجال في المجتمع . ويمكن تلخيص هذه المقـارنة بسهـولة من خلال النسبـة المئوية للـرجال الذين يتـوقع حصولهم على درجـات في هذا الاختبـار أسوأ من 55 وللحصول على هـذه النسبة نحسب المساحة تحت منحنى الكثافة للتوزيع $N(50, 100)$ الواقعة إلى اليسار من النقطة 55 وبمعايرة الدرجة 55 تأخذ القيمة :

$$Z = \frac{55 - 50}{10} = 0.5$$



شكل (٥ ـ ٦) : التـوزيع N (50, 100) لدرجات اختبـار المهارة الشفهية ، والمساحـة المظللة هي احتمال الحصول على درجة أقل من 55 .

والمساحة المطلوبة هي إذا المساحة الواقعة إلى اليسار من النقطة 0.5 تحت منحنى الكثافـة الطبيعي المعياري والمبينـة في الشكل (٥ ـ ٧) . وهي تسـاوي من الجدول ١ في



شكل (٥ ـ ٧) درجة الاختبار بعد معايرتها.

الملحق 0.6915 . وهكذا نستنتج أن 69% من المجتمع يتوقع حصولهم على درجة أسوأ، و 31% من المجتمع يتوقع حصولهم على درجة أفضل وهذا يحدد بوضوح موقعه النسبي من الآخرين.

ولو فرضنا أن درجة هذا الشاب كانت 40 في اختبار لقياس المهارات الحسابية. وهذا الاختبار مصمم بدوره بحيث يكون توزيع الدرجات الناتجة عنه $N(50, 100)$. وبمعايرة هذه الدرجة نجد أنها تصبح في سلم القياس المعياري:

$$Z = \frac{40 - 50}{10} = -1$$

وبالعودة إلى جدول التوزيع الطبيعي المعياري نجد أن

$$F(Z) = F(-1) = 1 - F(1) = 1 - 0.8413 = 0.1587$$

وهكذا نتوقع أن ينال 16% فقط من المجتمع درجات أسوأ، وأن ينال 84% درجات أفضل.

وبصورة عامة، تسمى معايرة متغير طبيعي X توزيعه $N(\mu, \sigma^2)$ ، أي التحويل من X إلى المتغير الطبيعي Z وفق العلاقة

$$Z = \frac{X - \mu}{\sigma}$$

تعبيرا عن قيمة المتغير X وفق سلم القيـاس المعياري. وهو سلـم قياس يعتبر μ مبدأ للقياسـات، ويعتبر الانحراف المعياري σ وحدة قياس. وعنـدما لا نهتم بقيمة X لذاتها بل بموقع X النسبي من المتوسط μ ، فإن القيمة Z توضح لنا بالضبط هذا الموقع النسبي ومنطوق العبارة الجبرية $X - \mu = Z\sigma$ ، هو أن موقع X يحيد عن النقطة μ بمقدار Z مرة الانحراف المعياري.

مثال (٥ ـ ٣)

إذا كانت درجات حـاصل الذكاء تتوزع طبيعيـا بمتوسط يساوي 100 وانحراف معياري يساوي 15 ، فما نسبة الناس ذوي درجة ذكاء:

ا ـ فوق 125 ، تحت 80 ، بين 70 و 130؟

الحل

لنرمـز لدرجـة حـاصـل الذكـاء بـ X ، فلدينـا بالفـرض أن توزيـع X هو $N(100, 152)$. والمطلوب

أ ـ

$$P(X > 125) = 1 - F\left(\frac{125 - \mu}{\sigma}\right) = 1 - F\left(\frac{125 - 100}{15}\right)$$
$$= 1 - F(1.67) = 1 - 0.9525 = 0.0475$$

والنسبة المطلوبة هي 4.75% .

ب ـ

$$P(X < 80) = F\left(\frac{80 - \mu}{\sigma}\right) = F\left(\frac{80 - 100}{15}\right)$$
$$= F(-1.33) = 1 - F(1.33) = 1 - 0.9082 = 0.0918$$

والنسبة المطلوبة هي 9.18% .

جـ ـ

$$P(70 < X < 130) = F\left(\frac{130 - \mu}{\sigma}\right) - F\left(\frac{70 - \mu}{\sigma}\right)$$
$$= F\left(\frac{130 - 100}{15}\right) - F\left(\frac{70 - 100}{15}\right) = F(2) - F(-2)$$
$$= 2F(2) - 1 = 2(0.9772) - 1 = 1.9544 - 1 = 0.9544$$

والنسبة المطلوبة هي 95.44%.

وكثيرا ما نستخدم علاقة المعايرة، $Z=\frac{X-\mu}{\sigma}$ ، بطريقة عكسية. فنحن نعرف أو نحدد سلفا قيمة Z ، أي القياس المطلوب على السلم المعياري، ونريد القياس المقابل له على السلم الأصلي (قبل المعايرة). فلنفرض، مثلا، أن لدى مدير شركة وظيفة شاغرة، وهو لا يقبل مرشحين لهذه الوظيفة إلا إذا كانوا في مهاراتهم الحسابية من الربع الأعلى في المجتمع. ولترجمة رغبته هذه بدلالة الدرجة الدنيا التي ينبغي أن ينالها المرشح في اختبار المهارات الحسابية، نقوم بما يلي، مفترضين أن درجات الاختبار تتبع التوزيع $N(50, 100)$. نحدد من عبارة «المرشح من الربع الأعلى في المجتمع في مهاراته الحسابية» قيمة Z ، وذلك لأن هذه العبارة مكافئة للمعادلة $F(Z) = 0.75$ ، ومن جدول التوزيع الطبيعي المعياري نجد، باستخدام الاستيفاء، أن

$$Z = 0.67 + \left(\frac{14}{31}\right)(0.01) = 0.6745$$

وبالتالي

$$X = \mu + Z\sigma = 50 + 10\,(0.6745) = 56.745$$

وبالتدوير إلى أقرب عدد صحيح، نستنتج أن الدرجة المطلوبة هي 57 وهكذا لا يقبل طلب متقدم لهذه الوظيفة إلا إذا كانت درجته في اختبار المهارات الحسابية 57 أو أكثر.

مثال (٥ ـ ٤)

بالإشارة إلى المثال (٥ ـ ٣) وتوزيع درجات حاصل الذكاء. لنفرض أن الحكومة تقدم تعليما خاصا للخمسة في المائة الأدنى في حاصل ذكائهم. وتقدم تعليما جامعيا للسبعة في المائة الأعلى في حاصل ذكائهم. أوجد القيم المعيارية Z المقابلة لهذه النسب ثم استنتج الحدود الفاصلة في درجات حاصل الذكاء لأولئك الذين يتطلبون تعليما خاصا، ولأولئك الذين يدخلون الجامعات.

الحل

لنفرض أن القيمة المعيارية المقابلة لنسبة جماعة التعليم الخاص هي a ، والمقابلة لنسبة جماعة التعليم الجامعي هي b فعندئذ:

$P(Z \leq a) = 0.05, F(a) = 0.05, F(-a) = 0.95, -a = 1.645, a = -1.645.$

$P(Z > b) = 0.07; 1 - F(b) = 0.07, F(b) = 0.93$

$b = 1.47 + 8(0.01)/14 = 1.47 + 0.0057 = 1.4757$

ويكون الحد الأعلى لدرجة حاصل الذكـاء لأولئك الذين يتطلبون تعليما خاصا ، مقربا إلى أقرب عدد صحيح هو:

$$X = \mu + a\sigma = 100 + 15(-1.645) = 75$$

والحد الأدنى لدرجة حاصل الذكاء لأولئك الذين يدخلون الجامعات، مقربا إلى أقرب عدد صحيح، هو:

$$X = \mu + b\sigma = 100 + 15(1.4757) = 122$$

مثال (٥ ـ ٥)

إذا كـــان X متغيرا طبيعيـــا متوسطـــه $\mu = 56$ وانحرافـه المعيـاري $\sigma = 3$ ، فاحسب $P(X \leq 60.5)$ ،$P(X > 65)$ ،$P(53 < X < 59)$.

الحل

$$P(X \leq 60.5) = P\left(\frac{X-\mu}{\sigma} \leq \frac{60.5-\mu}{\sigma}\right) = P\left(Z \leq \frac{60.5-56}{3}\right)$$
$$= F\left(\frac{60.5-56}{3}\right) = F(1.5) = 0.9332 .$$

و

$$P(X > 65) = P\left(\frac{X-\mu}{\sigma} > \frac{65-\mu}{\sigma}\right) = P\left(Z > \frac{65-56}{3}\right)$$
$$= 1 - F\left(\frac{65-56}{3}\right) = 1 - F(3) = 1 - 0.9987 = 0.0013$$

و

$$P(53 < X < 59) = F\left(\frac{59-56}{3}\right) - F\left(\frac{53-56}{3}\right)$$
$$= F(1) - F(-1) = 2F(1) - 1$$
$$= 2(0.8413) - 1 = 1.6826 - 1 = 0.6826 .$$

مثال (٥ ـ ٦)

ليكن X متغيرا عشوائيـا يتوزع وفق التـوزيع الطبيعي بمتوسـط يساوي 2 وتباين يساوي 16 . والمطلوب حساب احتمالات الحوادث العددية التالية :

$P(X < 3)$ ،$P(X > 1)$ ،$P(-1 < X < 35)$.

الحل

$$P(X<3)=F\left(\frac{3-\mu}{\sigma}\right)=F\left(\frac{3-2}{4}\right)$$
$$=F(0.25)=0.5987\ .$$
$$P(X>1)=1-F\left(\frac{1-\mu}{\sigma}\right)=1-F\left(\frac{1-2}{4}\right)$$
$$=1-F(-0.25)=F(0.25)=0.5987$$
$$P(-1<X<3.5)=F\left(\frac{3.5-\mu}{\sigma}\right)-F\left(\frac{-1-\mu}{\sigma}\right)$$
$$=F\left(\frac{3.5-2}{4}\right)-F\left(\frac{-1-2}{4}\right)$$
$$=F(0.375)-F(-0.75)$$
$$=F(0.375)-[1-F(0.75)]$$
$$=F(0.375)+F(0.75)-1$$

ولحساب $F(0.375)$ نأخذ منتصف الطريق بين $F(0.37)$ و $F(0.38)$ ، أي منتصف الطريق بين 0.6443 و 0.648 وهو إلى أربعة أرقام عشرية 0.6462 . وهكذا يكون

$$P(-1<X<3.5)=0.6462+0.7734-1=0.4196$$

مثال (٥ ـ ٧)

في عملية تعبئة آلية لعبوات السكر، من المفترض أن تضع الآلة في كل عبوة 2 كغ من السكر. وبالطبع يتغير ما تضعه الآلة من عبوة إلى أخرى بشكل عشوائي. إذا افترضنا أن ما تضعه الآلة بالفعل هو متغير $N(\mu, \sigma^2)$.

أ ـ تشير السجلات السابقة للإنتاج إلى أن $\sigma = 0.2$ ، وإلى أن احتمال أن تتضمن عبوة أقل من 2 كغ هو 0.01 . أوجد قيمة μ التي تعمل الآلة وفقا لها. (أي القيمة المتوسطة لما تضعه هذه الآلة في العبوة الواحدة على المدى الطويل.)

ب ـ إذا قمنا بعملية تحسين لعمل الآلة تتوخى تخفيض σ (أي انتاج عبوات أكثر تجانسا من حيث الوزن) مع بقاء μ كما هو . كم يجب أن تكون قيمة σ بحيث نطمئن إلى أن احتمال عبوة بأقل مما ينبغي من السكر هو 0.001 ؟

الحل

أ ـ لنرمز بـ X لوزن السكر الفعلي في العبوة . والمطلوب هو حساب μ علما أن

$$P(X < 2) = 0.01$$

و $\sigma = 0.02$. ولكن

$$P(X < 2) = F\left(\frac{2-\mu}{0.02}\right) = 0.01$$

أو

$$F\left(\frac{\mu-2}{0.02}\right) = 0.99$$

ومن الجدول نجد أن :

$$\frac{\mu-2}{0.02} = 2.33$$

أو

$$\mu = 0.02\,(2.33) + 2 = 2.047$$

ب ـ إذا اشتغلت الآلة وفقا لـ $\mu = 2.047$ فعندئذ يكون X متغيرا $N(2.047, \sigma^2)$.

ونريد قيمة σ بحيث يكون :

$$P(X < 2) = 0.001$$

ولكن الآن :

$$P(X < 2) = F\left(\frac{2-2.047}{\sigma}\right)$$

إذا نريد σ بحيث يكون

$$F\left(\frac{-0.047}{\sigma}\right) = 0.001$$

أو

$$F\left(\frac{0.047}{\sigma}\right) = 0.999$$

ومن الجدول نجد :

$$\frac{0.047}{\sigma} = 3.09$$

أي أن

$$\sigma = \frac{0.047}{3.09} = 0.015$$

مثال (٥ ـ ٨)

مفترضا أن طول الذكر البالغ X، مقاسا بالسنتمتر، هو متغير $N(175, 56.25)$. كيف يحدد مهندس ارتفاع أبواب الغرف في فيلا يقوم بتصميمها بحيث لا يضطر أكثر من 2% من الرجال إلى طأطأة رؤوسهم عند الدخول أو الخروج؟

الحل

لنفرض أن ارتفاع الباب a سم فيكون المطلوب تحديد قيمة a بحيث يكون :

$$P(X > a) \leq 0.02$$

ولكن

$$P(X > a) = 1 - P(X < a) = 1 - F\left(\frac{a - 175}{7.5}\right) \leq 0.02$$

وبالتالي يكون

$$F\left(\frac{a - 175}{7.5}\right) \geq 0.98$$

ومن الجدول نجد أن :

$$\frac{a - 175}{7.5} \geq 2.057$$

وهكذا يكون :

$$a \geq 175 + 2.057\,(7.5) = 190.43$$

أي أن ارتفاع الباب ينبغي أن يكون 190.5 سم على الأقل .

تمارين (٥ ـ ٢)

١) باستخدام جدول التوزيع الطبيعي أحسب الاحتمالات التالية، حيث Z المتغير الطبيعي المعياري $N(0, 1)$.

$P(Z \leq 1.2)$، $P(-0.9 < Z < 0)$، $P(0.3 < Z < 1.56)$، $P(|Z| < 0.2)$، $P(-1.3 < Z < 1.74)$، $P(Z > -0.75)$، $P(Z \leq -0.32)$.

٢) أوجد المساحة تحت منحنى كثافة التوزيع الطبيعي المعياري الواقعة :

أ ـ إلى اليسار من 1 ،

ب ـ إلى اليسار من 2 ،

جـ ـ بين 1 و 2 ،

د ـ إلى اليمين من 0.5 – ،

هـ ـ إلى اليسار من 1– ،

و ـ بين 1– و 1+ .

٣) أوجد العدد c بحيث يكون :

أ ـ $P(Z < c) = 0.8643$ ،

ب ـ $P(Z < c) = 0.2266$ ،

جـ ـ $P(Z \geq -c) = 0.6554$ ،

د ـ $P(Z < c) = 0.05$

هـ ـ $P(-c < Z < c) = 0.90$ ،

و ـ $P(-c < Z < c) = 0.95$ ،

ز ـ $P(-c < Z < c) = 0.99$.

٤) إذا رمزنا بـ Z_α لقيمة المتغير الطبيعي المعياري Z التي تحصر إلى اليمين منها مساحة تساوي α ، فاحسب $Z_{0.005}$ ، $Z_{0.025}$، $Z_{0.05}$، $Z_{0.02}$، $Z_{0.01}$، $Z_{0.10}$.

٥) متغير عشوائي X يتبع التوزيع الطبيعي $N(16, 7)$ ، احسب $P(|X - 16| > 3)$.

٦) متغير عشوائي X يتبع التوزيع $N(50, 25)$ ، احسب :
$P(X > 62)$، $P(|X - 50| < 8)$، $P(X = 60)$، $P(|X - 40| > 5)$.

٧) تتوزع معدلات مجتمع كبير من طلبة الكليات تقريبا وفق التوزيع $N(2.4, 0.64)$. ما نسبة الطلاب الذين تتجاوز معدلاتهم 3.0 ؟ (المعدل التام هو 4) .

٨) بالإشارة إلى المسألة السابقة إذا شطبت أسماء الطلاب الذين تقل معدلاتهم عن 1.9 فكم ستبلغ نسبة الأسماء المشطوبة؟

٩) متغير عشوائي X يتبع التوزيع الطبيعي. إذا كان $E(X^2) = 68$ و $P(X < 10) = 0.8413$ فاحسب μ و σ^2.

١٠) يتوزع عمر نوع من الغسالات مقدرا بالسنوات وفق التوزيع الطبيعي $N(3.1, 1.2)$. إذا كانت الغسالات مكفولة لمدة سنة، فما هي نسبة الغسالات المباعة التي سيضطر المصنع إلى استبدالها بغسالة جديدة؟

١١) وجدنا أن الفترة الزمنية الضرورية لإتمام اختبار ذكاء مخصص لطلبة الكليات تتوزع وفق التوزيع الطبيعي بمتوسط 70 دقيقة وانحراف معياري يساوي 12 دقيقة. كيف يجب تحديد زمن الاختبار إذا أردنا إتاحة وقت كاف لإتمام الاختبار لـ 90% من الطلاب المتقدمين؟

١٢) نظمت آلة لتقديم شراب مرطب بحيث تضع، في المتوسط، μ أونزة في الكأس الواحدة. إذا كان ما تضعه بالفعل في الكأس الواحدة متغيرا طبيعيا بانحراف معياري $\sigma = 0.3$ أونزة. فما القيمة التي ينبغي تحديدها لـ μ بحيث تفيض الكؤوس ذات السعة 8 أونزة بنسبة 1% فقط؟

١٣) وزن بيضة الدجاج بالغرام يتوزع وفق التوزيع الطبيعي $N(60, 225)$. وتصنّف البيضة «صغيرة» إذا قل وزنها عن 45 غراما، إذا رغبت أن يصنف باقي البيض بالتساوي بين عادي وكبير، إقترح الوزن الذي يفصل بين هذين الصنفين مقربا إلى أقرب غرام.

١٤) تتوزع أوزان قوالب الصابون في مصنع طبيعيا. وفي الأسبوع الماضي كان وزن $6\frac{2}{3}\%$ من القوالب المصنوعة أقل من 90.5 غراما بينما زاد وزن 4% من القوالب على 100.25 غراما. والمطلوب:

أ - أوجد متوسط وتباين توزيع وزن القالب، والنسبة المئوية للقوالب التي يتوقع أن تزن أقل من 88 غراما.

ب ـ إذا خفضنا تباين الوزن بنسبة الثلث فما هي النسبة المئوية من إنتاج الأسبوع القادم التي تتوقع أن يقل وزنها عن 88 غراما. مفترضا أن المتوسط لم يتغير؟

١٥) يقدر أن 1400 راكبا ممن يبدلون قطارهم في محطة معينة يهدفون بصورة منتظمة إلى اللحاق بقطار الخامسة والنصف مساء، وأن 50 راكبا يصلون قبل الساعة الخامسة وعشرين دقيقة مساء، موعد فتح البوابة الخاصة بهذا القطار، وأن 70 راكبا يفوتهم القطار عند التزامه التام بموعد المغادرة. مفترضا أن زمن وصول الراكب إلى المحطة متغير عشوائي يتبع التوزيع الطبيعي، أحسب متوسط هذا التوزيع وتباينه. ومن ثم قدر:

ا ـ موعد فتح البوابة بحيث لا يزيد عدد المنتظرين أمامها على عشرين راكبا.

ب ـ عدد المستبدلين الذين سيفوتهم القطار في يوم يغادر فيه (على غير المتوقع) قبل الوقت المحدد بدقيقتين.

١٦) يغادر رجل منزله كل صباح الساعة السابعة كي يصل إلى عمله في الساعة الثامنة. وقد وجد خلال فترة طويلة أنه يتأخر عن عمله بنسبة مرة في كل أربعين مرة. وبدأ يغادر المنزل في الساعة السادسة وخمس وخمسين دقيقة فوجد خلال فترة مماثلة أنه يتأخر مرة في كل مائة مرة. بفرض أن الزمن الذي تستغرقه الرحلة يتوزع طبيعيا كيف ينبغي أن يحدد موعد المغادرة بحيث لا يتأخر أكثر من مرة كل مائتي مرة؟

١٧) في كتاب معين يمكن اعتبار عدد الكلمات في الصفحة الواحدة متغيرا طبيعيا، على وجه التقريب، بمتوسط 800 كلمة وانحراف معياري 50 كلمة. إذا اخترت عشوائيا ثلاث صفحات فما احتمال ألا تتضمن أي منها ما بين 830 إلى 845 كلمة؟

١٨) في بلد معين، متوسط طول الذكر البالغ 170 سم بانحراف معياري 10 سم، ومتوسط طول الأنثى البالغة 160 سم بانحراف معياري 8 سم، وبالنسبة لكل

مـن الجنسين يعتبر التوزيـع الطبيعي نموذجا منـاسبا لوصف تغير الطول. بفرض أن الطول ليس من العـوامل التي تؤخـذ في الاعتبـار عنـد اختيار الزوجة أو الزوج. أحسب احتمـال أن زوجـا وزوجته اخترناهما عشـوائيا سيكـون كل منهما أطول من 164 سم.

١٩) في بستـان للبرتقال متـوسط وزن الثمـرة 19.3 أونزة بـانحراف معيـاري 2.3 أونزة. مفترضا أن وزن الثمرة متغير يتبع التوزيع الطبيعي، أحسب:

ا ـ نسبة الثمار التي يقل وزنها عن 18 أونزة.

ب ـ نسبة الثمار التي لا يقل وزنها عن 20 أونزة.

جـ ـ نسبة الثمار التي يتراوح وزنها بين 18.5 و 20.5 أونزة.

د ـ الوزن الذي سيقل عنه 15% من الثمار.

هـ ـ الوزن الذي سيزيد عليه 25% من الثمار.

٢٠) ملاحظة عدد كبير من السيارات عند نقطة محددة من طريق عام بينت لنا أن السرع تتوزع طبيعيا. إذا علمت أن سرعة 90% من السيارات تقل عن 124.3 كم/ سا، وأن سرعة 5% فقط من السيـــارات تقل عن 101 كم/ س. حـدد السرعة المتـوسطة μ والانحراف المعياري σ.

٢١) من المفترض أن يكون قطـر كريات معدنيـة تنتجها شركة صناعيـة مساويا 2 مم. ولكن الكريات ستكون مقبولـة إذا تراوحت أقطارها بين 1.90 مم و 2.10 مم. وقد لوحظ في دفعـة إنتاج كبيرة أن 2.5% منها مرفوض لأنـه أكبر مما يمكن التساهل فيه وأن 2.5% منها مرفوض لأنه أصغر مما يمكن التساهل فيه. حدد، بصورة تقريبية، ما ستصبحه نسبتا الرفض إذا غيرنا حدود التساهل إلى 1.95 مم و 2.15 مم.

٢٢) تتوزع درجات امتحان وفق التوزيع الطبيعي $N(50, 100)$، ونرغب في إعادة النظر في سلم الدرجات بحيث تكـون درجة النجاح 40 ونسبة الناجحين 70%، ودرجة

التفوق 70 ونسبة المتفوقين 20% . أحسب الدرجة الجديدة لمتقدم للامتحان كانت درجته الأصلية 60 .

٢٣) يمكن تصنيف البيض إلى عادي إذا كان الوزن أقل من 46 غراما، ومتوسط إذا كان الوزن بين 46 و 56 غراما، وكبير إذا كان الوزن أكبر من 56 غراما . لنفرض أن البيض الذي تضعه سلالة معينة من الدجاج يتوزع، من حيث وزن البيضة، وفق التوزيع الطبيعي $N(50, 25)$. أحسب نسبة كل صنف من الأصناف الثلاثة . وإذا كانت أسعار البيع للبيضة الواحدة من الأصناف الثلاثة هي، على الترتيب، 4 هللة، 5 هللة،6 هللة . وكانت كلفة الإنتاج 4 هللة لكل بيضة، فما الربح المتوقع للبيضة الواحدة؟

وبالنسبة لسلالة أخرى من الدجاج فإنها تضع بيضا يتبع، من حيث الوزن، التوزيع الطبيعي $N(52, 25)$. إلا أنه يستهلك أكثر من الطعام مما يرفع كلفة البيضة إلى 4.5 هللة . ما الربح المتوقع للبيضة الواحدة في هذه السلالة؟

٢٤) لنفرض أن مقاس الحذاء لذكر بالغ هو عدد صحيح k يرتبط بطول القدم، y ، مقاسا بالبوصة بالعبارة التالية : «حذاء مقاسه k سيكون مناسبا لقدم طولها يتراوح بين $5.5 + 0.5k$ و $6 + 0.5k$ ؛ حيث $k = 5, 6, \ldots, 14$ » . ويمكن اعتبار y ، طول قدم ذكر بالغ، متغيرا يتبع التوزيع الطبيعي $N(10.2, 1.21)$.

ا ـ ما النسبة من مجتمع الذكور البالغين التي تتطلب حذاء مقاسه أكبر من 14؟

ب ـ ما المقاس الأكثر تواترا وما نسبة أولئك الذين يطلبون هذا المقاس؟

٢٥) حدود التساهل في طول قطعة مصنعة هي 10.00 ± 0.05 مم . وتُفحص كل قطعة يجري إنتاجها لرؤية ما إذا كانت تحقق هذه الحدود أم لا . والتوزيع الاحتمالي لطول القطعة هو التوزيع الطبيعي بمتوسط يساوي 10.01 مم وانحراف معياري 0.04 مم . وكلفة إنتاج القطعة 10 ريالات . وجميع القطع التي لا يقع طولها ضمن حدود

التساهل تهمل وتعتبر خسارة للشركة المصنعة. ولتخفيض حجم الخسارة يمكن:

ا - إزالة الانحياز في عمل الآلة وجعل متوسط التوزيع $\mu = 10$ وذلك بكلفة إضافية قدرها 4 ريالات لكل قطعة.

ب - تخفيض الانحراف المعياري إلى 0.03 وذلك بكلفة إضافية قدرها ريالان لكل قطعة.

جـ - القيام بالإجرائين (ا) و (ب) معا لقاء كلفة إضافية 6 ريالات للقطعة الواحدة. إذا كنت تعمل في قسم الإحصاء في هذه الشركة فبأي الإجراءات الثلاثة المذكورة تنصح؟

٢٦) تقضي مواصفات الإنتاج لعبوات نوع معين من الحلويات أن وزن كل عبوة يجب أن يقع بين 140 غ و 160 غ. إذا كان وزن العبوة يتوزع طبيعيا بتباين يساوي 4غ^{٢}. كيف تحدد متوسط التوزيع الذي ينبغي أن تهدف إليه الشركة المنتجة ولماذا؟

٢٧) يستخدم أحد المصانع 2000 مصباح كهربائي للإضاءة. وعمر المصباح الكهربائي مقاسا بالساعات يتبع التوزيع الطبيعي $N(550, 2500)$. وحرصا على وجود عدد قليل من المصابيح المحترقة خلال أوقات الإنتاج يستبدل المصنع المصابيح جميعها كل فترة وبصورة دورية. كيف ينبغي تحديد طول فترة الاستبدال لكي لا يوجد في المصنع في أي وقت أكثر من 20 مصباحا محروقا؟

ومع نوع أفضل من المصابيح حيث يتوزع عمر المصباح وفق التوزيع الطبيعي $N(600, 1600)$ تتغير فترة الاستبدال إلى 500 ساعة، بين أن عدد المصابيح المحترقة في المصنع في أي وقت سينخفض عندئذ إلى حوالي 12 مصباحا.

٢٨) مبيعات بقال من سلعة معينة كل أسبوعين هي متغير يتبع التوزيع الطبيعي بمتوسط يساوي 200 كغ وتباين يساوي 225 كغ. أوجد احتمال أن تكون مبيعاته من هذه السلعة خلال أسبوعين أقل من 185 كغ، وعندما يطلب مزيدا من هذه

السلعة تأخذ عملية تسليم البضاعة المطلوبة فترة أسبوعين. حدد إلى أقرب كيلوغرام المخزون الذي ينبغي تأمينه من هذه السلعة عند إعادة طلبها بحيث يكون البقال مطمئنا باحتمال 0.95 إلى أن هذه السلعة لن تنفذ قبل وصول الطلب.

(٥ ـ ٤) خواص التوزيع الطبيعي وبعض التطبيقات*

اصطلحنا على كتابة $N(\mu, \sigma^2)$ لتعني توزيعا طبيعيا بمتوسط يساوي μ وتباين يساوي σ^2 . وهكذا نكتب، على سبيل المثال: X متغير $N(8, 4)$ لتعني أن X متغير يتبع التوزيع الطبيعي بمتوسط يساوي 8 وتباين يساوي 4 . وفيما يلي بعض خواص التوزيع الطبيعي:

١ ـ ليكن X و Y متغيرين مستقلين $N(\mu_1, \sigma_1^2)$ و $N(\mu_2, \sigma_2^2)$ ، على الترتيب. فعندئذ يكون مجموعهما $X + Y$ ، ولنرمز له بـ U ، متغيرا $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. أي متغيرا طبيعيا أيضا بمتوسط يساوي مجموع المتوسطين وتباين يساوي مجموع التباينين.

٢ ـ وبصورة أعم إذا كان X و Y متغيرين مستقلين $N(\mu_1, \sigma_1^2)$ و $N(\mu_2, \sigma_2^2)$ ، على الترتيب فإن المتغير $U = aX + bY + c$ ، حيث a ، b ، c أية أعداد حقيقية، هو بدوره متغير طبيعي متوسطه، حسب خواص التوقع:

$$E(U) = E(a\,X + b\,Y + c) = a\,E(X) + b\,E(Y) + c$$
$$= a\,\mu_1 + b\,\mu_2 + c$$

وتباينه حسب خواص التباين:

$$V(U) = V(a\,X + b\,Y + c) = V(a\,X + b\,Y) = V(a\,X) + V(b\,Y)$$
$$= a^2\,V(X) + b^2\,V(Y)$$
$$= a^2\,\sigma_1^2 + b^2\,\sigma_2^2$$

* للقراءة فقط.

ونكتب باختصار:

إذا كان X و Y متغيرين مستقلين $N(\mu_1, \sigma_1^2)$ ، $N(\mu_2, \sigma_2^2)$ وكانت a ، b ، c أية أعداد ثابتة فإن $U = aX + bY + c$ يكون متغيرا

$$N\left(a\,\mu_1 + b\,\mu_2 + c,\; a^2\,\sigma_1^2 + b\,\sigma_2^2\right)$$

وعلى سبيل المثال إذا كان X متغيرا $N(15, 2)$ و Y متغيرا $N(-7, 4)$ فإن $U = 2X - 3Y + 1$ وهو متغير طبيعي متوسطه يساوي

$$2(15) - 3(-7) + 1 = 52$$

وتباينه

$$2^2(2) + (-3)^2(4) = 44$$

أي أن U متغير $N(52, 44)$.

٣ ـ ويمكن بوضوح تعميم الخاصة ٢ إلى أكثر من متغيرين، لتصبح في الحالة الخاصة التالية، وهي في حد ذاتها بالغة الأهمية، كما يلي:

إذا كانت $X_1, X_2, \ldots, X_n$ متغيرات مستقلة وكل منها $N(\mu, \sigma^2)$ ، [أي إذا كانت $X_1, X_2, \ldots, X_n$ عينة عشوائية من $N(\mu, \sigma^2)$] فإن:

$$\sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$$ يكون متغيرا $N(n\mu, n\sigma^2)$ ،

ويكون

$$\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$$ متغيرا $N\left(\mu, \frac{\sigma^2}{n}\right)$.

وتجدر ملاحظة أنه بالرغم من أن المتغير الطبيعي يتحول بين $-\infty$ و $+\infty$ ، إلا أنه يمكن استخدامه استخداما مقبولا تماما لوصف متغير، X ، موجب بطبيعته. وذلك

شريطة أن يكـون $P(X \leq 0)$ عـددا صغيرا جدا يمكن إهمالـه . أي أننا نتجـاوز المقولـة الدقيقة بأن $P(X \leq 0) = 0$ ، وتعني استحالة أن يكون X سالبا إلى مقولة ، تقريبية وعملية في آن واحد ، تكتفـي بالتأكيـد على أن احتمال أن يكون X سالبا هـو احتمال قريب من الصفر. وبما أن

$$p(X \leq 0) = F\left(\frac{0-\mu}{\sigma}\right) = F\left(\frac{-\mu}{\sigma}\right)$$

وأن σ موجب ، فإن $P(X \leq 0)$ سيكون مهملا إذا كان μ كبيرا بالمقارنة مع σ .

وعلى سبيل المثال ، إذا كان $\mu = 4.5\sigma$ فإن $P(X \leq 0)$ يكون أقل من 0.000005 ، وهو صغير إلى الحد الذي يجعله غير ذي بال في التطبيقات العملية .

مثال (٥ ـ ٩)

إذا كانت X ، Y ، T متغيرات مستقلة $N(2, 1)$ ، $N(3, 2)$ ، $N(4, 3)$ ، على الترتيب ، فاحسب :

أ ـ $P(1 < X < 3)$ ،

ب ـ $P(X \leq Y)$ ،

جـ ـ $P(3X - 2Y > 1)$ ،

د ـ $P(X + Y < 2T - 4)$ ،

الحـــل

أ ـ
$$P(1 < X < 3) = F(3-2) - F(1-2) = F(1) - F(-1)$$
$$= 2F(1) - 1 = 0.6826$$

ب ـ
$$P(X \leq Y) = P(X - Y \leq 0)$$

ولكن $X - Y$ متغير $N(-1, 3)$ وفق الخاصة ٢ . وبالتالي يكون

$$P(X \leq Y) = P(X - Y \leq 0) = F\left(\frac{0-(-1)}{\sqrt{3}}\right) = F(0.577) = 0.718$$

جـ ـ وفق الخاصة ٢ نجد أن $3X - 2Y$ متغير $N(0, 17)$ وهكذا نجد :

$$P(3X - 2Y > 1) = 1 - P(3X - 2Y \leq 1) = 1 - F\left(\frac{1-0}{\sqrt{17}}\right)$$

$$= 1 - F(0.243) = 0.404$$

د ـ ووفق الخاصة ٣ يكون $X + 2Y - 2T$ متغيرا $N(-3, 15)$ ، وبالتالي :

$$P(X + Y \leq 2T - 4) = P(X + Y - 2T \leq -4) = F\left(\frac{-4-(-3)}{\sqrt{15}}\right)$$

$$= 1 - F\left(\frac{1}{\sqrt{15}}\right) = 1 - F(0.258) = 0.398 .$$

مثال (٥ ـ ١٠)

يتم إنتاج مسامير البرشام التي تستخدم لبرشمة صفيحة معدنية بطريقة تسمح لنا بوصف قطر المسمار X كمتغير $N(3; 0.04)$. وبطريقة مستقلة يجري إنتاج صفائح معدنية ذات ثقوب دائرية يمكن اعتبار قطر الثقب Y متغيرا $N(3.2, 0.01)$. (القياس في الحالتين بالسنتمتر) .

ا ـ ما هو احتمال أن يناسب المسمار ثقب الصفيحة؟

ب ـ إذا اخترنا أربعة أزواج (مسمار ـ صفيحة) فما هو احتمال أن يكون زوجان منهما ، على الأقل ، متناسبين؟

الحل

ا ـ X و Y متغيران طبيعيان مستقلان . واحتمال تناسب المسمار مع الثقب هو:

$$P(X < Y) = P(X - Y < 0)$$

ولكن $X - Y$ متغير $N(-0.2, 0.05)$ ، وبالتالي :

$$P(X - Y < 0) = F\left(\frac{0-(-0.2)}{\sqrt{0.05}}\right) = F(0.894) = 0.814$$

ب ـ يمكننا اعتبار إنتاج مسمار وصفيحة تكرارا لتجربة ثنائية احتمال النجاح فيها $p = 0.814$ ، $n = 4$ ، وإذا رمزنا بـ U لعدد الأزواج المتناسبة ، يصبح المطلوب :

$$P(U \geq 2) = 1 - P(U=0) - P(U=1)$$
$$= 1 - (0.186)^4 - 4\,(0.814)\,(0.186)^3 = 0.978\,.$$

مثال (٥ ـ ١١)

$$P(|\bar{X} - \mu| > 2) = P(\bar{X} - \mu > 2) + P(\bar{X} - \mu < -2) \leq 0.025$$

أخذنا عينة عشوائية حجمها n من مجتمع طبيعي فيه $\mu = 10$ و $\sigma = 20$ ، ما هي أصغر قيمة ممكنة لـ n بحيث لا يزيد عن 0.025 احتمال أن يتجاوز الفرق بين متوسطي العينة والمجتمع المقدار 2 ؟

الحل

ليكن $\bar{X}$ متوسط العينة. نعلم من الخاصة ٣ أن $\bar{X}$ متغير $N\left(10, \frac{400}{n}\right)$. والمطلوب تحديد حجم العينة n بحيث يكون $P(|\bar{X} - \mu| > 2) \leq 0.025$ ، ولكن الحادثة $|\bar{X} - \mu| > 2$ تعني إما $\bar{X} - \mu > 2$ أو $\bar{X} - \mu < -2$ ، وبالتالي :

$$P(|\bar{X} - \mu| > 2) = P(\bar{X} - \mu > 2) + P(\bar{X} - \mu < -2) \leq 0.025$$

أو

$$P\left(\frac{\bar{X} - \mu}{20/\sqrt{n}} > \frac{2}{20/\sqrt{n}}\right) + P\left(\frac{\bar{X} - \mu}{20/\sqrt{n}} < \frac{-2}{20/\sqrt{n}}\right) \leq 0.025$$

أو

$$1 - F\left(\frac{2\sqrt{n}}{20}\right) + F\left(\frac{-2\sqrt{n}}{20}\right) \leq 0.025$$

أو

$$2 - 2F\left(\frac{\sqrt{n}}{10}\right) \leq 0.025$$

$$F\left(\frac{\sqrt{n}}{10}\right) \geq 0.9875$$

ونجد من الجدول أن

$$\frac{\sqrt{n}}{10} \geq 2.24$$

$$\sqrt{n} \geq 22.4 \Leftrightarrow n \geq 501.76$$

أي أن حجم العينة ينبغي ألا يقل عن 502 .

مثال (٥ ـ ١٢)

ليكن X متغيرا عشوائيا يتبع التوزيع الطبيعي $N(100, 25)$. أحسب $P(|\bar{X} - 100| > 1)$ ، إذا كان :

ا ـ $\bar{X}$ متوسط عينة حجمها $n = 25$ ،

ب ـ $\bar{X}$ متوسط عينة حجمها $n = 100$.

الحل

ا ـ بالاستناد إلى الخاصة ٣ نعلم أن $\bar{X}$ متغير يتبع التوزيع الطبيعي $N(100, 1)$ ويكون

$$P(|\bar{X} - 100| > 1) = P(\bar{X} - 100 > 1) + P(\bar{X} - 100 < -1)$$
$$= 1 - F(1) + F(-1)$$
$$= 1 - F(1) + [1 - F(1)] = 2 - 2F(1)$$
$$= 2 - 2 \times 0.8413 = 0.3174$$

ب ـ $\bar{X}$ يتبع الآن التوزيع الطبيعي $N(100, 0.25)$ ومنه :

$$P(|\bar{X} - 100| > 1) = P\left(\frac{\bar{X} - 100}{0.5} > \frac{1}{0.5}\right) + P\left(\frac{\bar{X} - 100}{0.5} < \frac{-1}{0.5}\right)$$
$$= 1 - F(2) + F(-2)$$
$$= 2[1 - F(2)] = 0.0456$$

تمارين (٥ ـ ٣)

١) في المثال (٥ ـ ١٢) ، كم يجب أن يكون حجم العينة n ليصبح :

ا ـ $P(|\bar{X} - 100| > 0.5) \le 0.01$ ،

ب ـ $P(|\bar{X} - 100| > 0.5) \le 0.001$.

٢) إذا افترضنا أن الدرجات في امتحان عام تتوزع ، على وجه التقريب ، وفق التوزيع الطبيعي $N\,(72,\ 100)$ ففي مجموعة عشوائية تتضمن مائة طالب ممن أدوا هذا الامتحان ، ما احتمال أن يختلف متوسط درجاتهم عن 72 بأكثر من 3 درجات؟

٣) ما أصغر حجم عينة ينبغي أخذها من مجتمع طبيعي فيه $\mu = 10$ و $\sigma = 20$ ، كي لا يزيد احتمال تجاوز متوسط العينة لضعف متوسط المجتمع عن 0.025 ؟

٤) إذا كان X ، Y و Z ثلاثة متغيرات مستقلة وتوزيعاتها ، على الترتيب ، $N(2, 2)$ ، $N(3, 3)$ ، $N(4, 4)$ ، فاحسب :

أ ـ $P(1 \le X \le 4)$ ،

ب ـ $P(X - 2 \le 4)$ ،

جـ ـ $P(2X + Y \ge 5)$ ،

د ـ $P(Z + 2 \le 4X - Y \le + 3)$ ،

هـ ـ $P(X \ge Y, Z - 3 > 0)$.

٥) يتوزع المتغيران المستقلان X و Y وفق $N(\mu, \sigma^2)$ و $N(2\mu, 2\sigma^2)$ ، على الترتيب .

ا ـ إذا كان $\sigma = 3$ و $P(X + 2Y \le 10)$ فاحسب μ .

ب ـ إذا كان $\mu = 0$ و $P(4X - Y < 3) = 0.4$ فاحسب σ .

جـ ـ إذا كان $P(|2X - Y| > 10) = 0.05$ و $P(Y \le s) = 0.9$ فاحسب μ و σ .

٦) يتوزع طول نصف قطر دولاب صغير ينتجه مصنع معين وفق التوزيع الطبيعي $N(1, 0.0001)$ (القياس بالسنتيمتر) . ويتم إنتاج الدواليب بصورة مستقلة ثم تجمع عقب ظهورها في خط الإنتاج أزواجا . ونعتبر أن الزوج من الدواليب مُرض إذا اختلف نصفا القطرين للدولابين بأقل من 0.03 سم .

ا _ ما نسبة الأزواج المرضية من الدواليب ؟
ب _ من بين خمسة أزواج ما احتمال أن يكون أحدها على الأقل غير مُرض ؟
جـ _ إلى أي حد ينبغي تخفيض الانحراف المعياري لطريقة الإنتاج كي تصبح نسبة الأزواج المرضية 99% ؟

٧) وجد طبيب يعمل في عيادة أن الأوقات التي تستغرقها استشارات المرضى مستقلة بعضها عن بعض، وتتبع التوزيع الطبيعي بمتوسط يساوي 5 دقائق وانحراف معياري 1.5 دقيقة. ويقابل مرضاه، على التتالي، بدون فواصل زمنية بين مريضين، مبتدئا عمله الساعة العاشرة صباحا. ما الموعد الذي ينبغي للمريض العاشر أن يرتبه مع سيارة أجرة بحيث يطمئن باحتمال 99% أن السيارة سوف لا تنتظره ؟ وإذا كان الطبيب سيقابل 22 مريضا قبل انصرافه، فما احتمال مغادرته للعيادة قبل الساعة 12 ظهرا ؟

٨) عمر قطعة إلكترونية مقاسا بالساعات يتوزع وفق التوزيع الطبيعي، لنفرض أن 92.5% من هذه القطع يتجاوز عمرها 2160 ساعة و 3.92% يتجاوز عمرها 17040 ساعة .

ا _ أحسب متوسط التوزيع وانحرافه المعياري .
ب _ إذا أخذنا عينة من 100 قطعة فاحسب احتمال أن يكون متوسط العمر في العينة :

(i) أكبر من 10000 ساعة ،
(ii) أقل من 8000 ساعة .
(iii) واقعا بين 8000 و 10000 ساعة .

٩) الأجر الأسبوعي بالريال الذي تدفعه شركة إلى عمالها يتوزع تقريبا وفق التوزيع الطبيعي $N(200, 324)$.

أ ـ أحسب احتمال ألا يختلف متوسط الأجر الأسبوعي لعينة عشوائية من 9 عمال عن متوسط المجتمع 200 بأكثر من 12 ريالا.

ب ـ كم يجب أن يكون حجم العينة حتى لا يختلف متوسطها عن متوسط المجتمع بأكثر من ستة ريالات إلا بنسبة بسيطة لا تتجاوز 10% ؟

: ١٠) على مدير شركة أن يقابل 20 مرشحا لوظيفة. ويعلم من تجربته السابقة أن وقت المقابلة مقاسا بالدقيقة يتوزع وفق التوزيع الطبيعي $N\ (10,\ 9)$. ويبدأ مقابلاته الساعة التاسعة صباحا. في أي وقت ينبغي له أن يطلب فنجان القهوة ويرتاح لمدة ربع ساعة إذا أراد أن يكون مطمئنا باحتمال 99% إلى أنه قد انتهى في ذلك الوقت من مقابلة 50% من المرشحين؟ وما احتمال أن ينتهي من كل المقابلات عند الساعة الواحدة بعد الظهر؟

١١) يتوزع وزن أمتعة المسافر جوا وفق التوزيع الطبيعي بمتوسط 20 كغ وانحراف معياري 5 كغ . ويتسع نوع معين من الطائرات لـ 100 راكب. ما هو احتمال أن يتجاوز الوزن الكلي لأمتعة المسافرين 2150 كغ؟

١٢) عمر صمام كهربائي مقاس بالساعة يتبع التوزيع الطبيعي $N\,(200,\ \sigma^2)$. إذا اشترى شخص عشر صمامات وأراد باحتمال 0.95 ألا يقل متوسط عمر الصمامات العشرة عن 190 ساعة، فما هي أكبر قيمة يمكن أن يأخذها الانحراف المعياري σ؟

١٣) بالإشارة إلى التمرين رقم ٢٨ من مجموعة التمارين (٥ ـ ٢)، لنفرض أن خمس بقالات متجاورة متضامنة بالنسبة إلى توفير تلك السلعة للزبائن. وأن مبيعاتها خلال أسبوعين من تلك السلعة مستقلة بعضها عن بعض وأن كلا منها تتبع

التوزيع الطبيعي بمتوسطات هي 200 ، 240 ، 180 ، 260 ، و 320 كغ ، وتباينات هي ، على الترتيب ، 225 ، 240 ، 225 ، 265 ، 270 كغ . اكتب متوسط وتباين الطلب على السلعة خلال أسبوعين ، وحدد إلى ثلاثة أرقام معنوية المستوى الإجمالي لمخزونها من تلك السلعة الذي ينبغي توفره عند طلب بضاعة جديدة بحيث يكون احتمال عدم نفاذها 0.99 .

احسب احتمال أن يتجاوز مجموع مبيعات البقالات الخمس من تلك السلعة خلال عشرة أسابيع 6200 كغ .

١٤) مصنع مربيات يضع في كل عبوة ثماني علب من ثمانية أنواع مختلفة . والمفروض أن تزن كل علبة 50 غراما . ولكن عمليا يتبع وزن كل علبة التوزيع الطبيعي N (52, 1.21) ، وبصورة مستقلة من نوع إلى آخر.

أ ـ ما نسبة العلب التي تزن أقل من 50 غراما؟

ب ـ ما نسبة العبوات التي تقل عن 400 غراما؟

جـ ـ ما احتمال أن تزن واحدة أو أكثر من العلب ضمن عبوة أقل من 50 غراما ؟

د ـ كم ينبغي أن يكون الانحراف المعياري لوزن العلبة إذا أردنا لـ 99% من العبوات أن تزن أكثر من 400 غراما؟

١٥) أوزان الأشخاص الذين يستخدمون مصعدا معينا تتوزع وفق التوزيع الطبيعي بمتوسط يساوي 150 ليبرة وانحراف معياري 20 ليبرة . والحد الأعلى المسموح لحمولة المصعد هو 650 ليبرة .

ا ـ بصورة عشوائية ، يجتمع أربعة أشخاص في المصعد . ما هو احتمال تجاوز الحمولة القصوى؟

ب ـ بصورة عشوائية يوجد شخص واحد في المصعد ومعه أمتعة تزن ثلاثة أمثال

وزنه، ما هو احتمال تجاوز الحمولة القصوى؟
فسر أي اختلاف بين جوابيك في (ا) و (ب).

(٥ ـ ٥) نظرية النهاية المركزية

تعرض نظرية النهاية المركزية، وتحت شروط عامة جدا، أن كلا من مجموع ومتوسط عينة عشوائية، مسحوبة من مجتمع ما، يمتلك عند تكرار هذه العينات عددا كبيرا من المرات، توزيعا له، على وجه التقريب، شكل الجرس. وربما كان من الأفضل إيضاح هذه العبارة بمثال.

لنعتبر المجتمع المتولد عن قذف حجر نرد عددا كبيرا جدا من المرات. وقد رأينا توزيعه في المثال (٣ ـ ٦). لنسحب عينة من خمسة قياسات، $n = 5$، من المجتمع وذلك بقذف حجر النرد خمس مرات وتسجيل الملاحظات الخمس الناتجة. ثم نحسب مجموع هذه الملاحظات الخمس Σx_i ومتوسطها $\bar{x}$، ويبين الجدول (٥ ـ ١) نتائج تكرار هذه العملية مائتي مرة. كما يبين الشكل (٥ ـ ٨) المدرج التكراري للقيم المائتين لـ $\bar{x}$ (أو Σx_i). وتنبغي ملاحظة النتيجة المهمة التالية:

بالرغم من أن التوزيع الاحتمالي لـ X له شكل أفقي تماما، إلا أن المدرج التكراري لمائتين من قيم $\bar{X}$ (وهو يقدم صورة أولية عن شكل التوزيع الاحتمالي للمتغير $\bar{X}$ أو للمتغير ΣX_i) يتخذ شكلا مقببا قريبا من شكل الجرس، وكلما زاد حجم العينات المسحوبة عن خمسة اعتدل شكل المضلع التكراري ليقترب أكثر فأكثر من شكل التوزيع الطبيعي. وبعبارة أخرى، لو أننا أخذنا $n = 10$ في مثالنا، أي لو أننا قذفنا حجر النرد عشر مرات بدلا من خمس، ثم سجلنا نتائج مائتي عينة من هذا الحجم، ورسمنا المدرج التكراري للقيم المائتين لـ $\bar{X}$، فمن المتوقع الحصول على

شكل أكثر قربا من شكل الجرس. ولا بد من ملاحظة أنه للحصول على فكرة أدق عن شكل التـوزيع الإحتمالي لــ $\bar{x}$ نحتاج ، نظـريـا ، إلى عدد لا نهائي من العينـات ، أو لنقل ، بصـورة عمليـة ، إننا نحتـاج إلى عـدد من العينـات أكبر بكثير من المائتين التي تضمنتهـا التجربـة هنـا . ومع ذلك فإن الشكل الـذي تقدمـه العينـات المائتان كـاف

جدول (٥ ـ ١): مئتا عينة من مجتمع قذف حجر نرد

رقم العينة	قياسات العينة	Σx_i	$\bar{x}$	رقم العينة	قياسات العينة	Σx_i	$\bar{x}$
1	3,1,6,4,1	15	3.0	33	6,3,5,4,5	23	4.6
2	4,6,6,5,2	23	4.6	34	6,5,3,3,3	20	4.0
3	5,5,2,5,2	19	3.8	35	2,6,2,6,3	19	3.8
4	4,4,5,2,2	17	3.4	36	2,2,1,6,6	17	3.4
5	2,3,6,3,3	17	3.4	37	4,3,2,5,4	18	3.6
6	6,6,2,5,4	23	4.6	38	5,1,2,5,6	19	3.8
7	6,3,3,2,6	20	4.0	39	5,5,2,5,6	23	4.6
8	3,1,5,1,5	15	3.0	40	5,6,6,5,2	24	4.8
9	6,2,5,5,4	22	4.4	41	3,1,6,3,6	19	3.8
10	6,5,6,6,6	29	5.8	42	1,6,2,6,1	17	3.4
11	6,6,1,1,2	16	3.2	43	3,2,3,4,6	18	3.6
12	1,4,1,4,6	16	3.2	44	3,2,5,1,6	17	3.4
13	4,6,3,5,5	23	4.6	45	4,6,5,3,2	20	4.0
14	4,3,3,4,5	19	3.8	46	6,2,5,4,5	22	4.4
15	4,6,2,3,1	16	3.2	47	6,1,1,2,5	15	3.0
16	1,4,3,4,5	17	3.4	48	1,1,5,5,2	14	2.8
17	3,4,3,1,4	15	3.0	49	2,2,3,3,4	14	2.8
18	3,3,3,6,4	19	3.8	50	5,4,2,2,1	14	2.8
19	6,3,4,4,6	21	4.2	51	3,5,1,5,3	17	3.4
20	5,4,2,2,6	19	3.8	52	5,2,3,3,2	15	3.0
21	4,5,5,2,2	18	3.6	53	4,1,5,2,6	18	3.6
22	1,5,2,3,1	12	2.4	54	5,4,4,2,4	19	3.8
23	3,5,6,5,3	22	4.4	55	4,5,2,1,4	16	3.2
24	5,3,6,4,3	21	4.2	56	4,5,6,3,1	19	3.8
25	6,2,3,2,5	18	3.6	57	3,5,5,1,4	18	3.6
26	5,4,5,1,6	21	4.2	58	6,6,5,3,4	24	4.8
27	4,1,6,2,6	19	3.8	59	6,3,2,5,4	20	4.0
28	6,6,6,2,2	22	4.4	60	4,6,5,1,1	17	3.4
29	3,4,2,1,5	15	3.0	61	5,1,1,2,2	11	2.2
30	1,2,2,3,3	11	2.2	62	2,6,2,2,3	15	3.0
31	6,5,1,6,2	20	4.0	63	2,4,4,1,1	12	2.4
32	6,3,1,2,5	17	3.4	64	3,1,2,2,2	10	2.0

تابع جدول (٥ ـ ١)

رقم العينة	قياسات العينة	Σx_i	$\bar{x}$	رقم العينة	قياسات العينة	Σx_i	$\bar{x}$
65	3,4,1,1,6	15	3.0	107	5,2,5,1,1	14	2.8
66	6,2,5,5,6	24	4.8	108	3,3,4,1,2	13	2.6
67	3,1,1,4,6	15	3.0	109	3,1,4,3,3	14	2.8
68	3,2,6,5,4	20	4.0	110	5,2,6,1,2	16	3.2
69	6,4,1,5,3	19	3.8	111	1,2,6,3,1	13	2.6
70	3,2,2,6,4	17	3.4	112	4,6,2,2,1	15	3.0
71	5,4,1,2,2	14	2.8	113	4,4,4,1,4	17	3.4
72	1,4,2,4,5	16	3.2	114	3,3,6,3,2	17	3.4
73	1,6,1,5,2	15	3.0	115	2,1,5,4,6	18	3.6
74	3,1,1,4,4	13	2.6	116	6,6,4,2,4	22	4.4
75	1,5,6,5,4	21	4.2	117	3,2,2,1,4	12	2.4
76	4,1,6,6,5	22	4.4	118	3,2,2,4,3	14	2.8
77	2,4,6,4,5	21	4.2	119	5,3,1,1,4	14	2.8
78	6,2,2,6,1	17	3.4	120	6,1,3,3,4	17	3.4
79	5,1,2,4,1	13	2.6	121	3,3,6,3,1	16	3.2
80	6,1,6,1,6	20	4.0	122	5,2,2,2,3	14	2.8
81	6,5,5,5,1	22	4.4	123	3,2,6,1,1	13	2.6
82	5,3,3,1,6	18	3.6	124	5,1,6,5,5	22	4.4
83	3,6,4,5,4	22	4.4	125	5,1,2,6,5	19	3.8
84	3,4,4,2,3	16	3.2	126	2,3,6,3,3	17	3.4
85	2,5,6,1,4	18	3.6	127	4,3,2,1,5	15	3.0
86	2,1,2,2,1	8	1.6	128	4,5,5,1,3	18	3.6
87	2,4,3,3,5	17	3.4	129	6,3,4,5,1	19	3.8
88	1,2,2,6,5	16	3.2	130	1,6,2,2,1	12	2.4
89	4,3,5,3,3	18	3.6	131	3,1,1,2,5	12	2.4
90	4,6,1,1,2	14	2.8	132	5,4,1,2,5	17	3.4
91	4,2,1,1,2	10	2.0	133	3,2,6,6,2	19	3.8
92	3,3,4,4,2	16	3.2	134	3,4,5,5,3	20	4.0
93	4,1,4,5,4	18	3.6	135	3,5,5,5,4	22	4.4
94	4,1,2,6,3	16	3.2	136	6,2,5,5,1	19	3.8
95	1,1,6,1,5	14	2.8	137	2,3,2,4,2	13	2.6
96	3,2,5,1,5	16	3.2	138	6,1,4,1,5	17	3.4
97	5,2,4,6,6	23	4.6	139	5,6,1,6,5	23	4.6
98	3,3,6,5,1	18	3.6	140	2,2,6,2,6	18	3.6
99	4,4,5,2,6	21	4.2	141	1,3,2,4,3	13	2.6
100	4,2,4,4,2	16	3.2	142	6,4,4,5,5	24	4.8
101	4,5,5,2,1	17	3.4	143	3,1,6,2,4	16	3.2
102	2,5,5,3,2	17	3.4	144	2,1,1,6,2	12	2.4
103	2,3,3,1,5	14	2.8	145	4,4,1,5,5	19	3.8
104	1,5,2,3,2	13	2.6	146	2,4,5,1,2	14	2.8
105	3,4,2,2,3	14	2.8	147	5,1,3,2,3	14	2.8
106	5,3,2,3,4	17	3.4	148	3,2,2,5,6	18	3.6

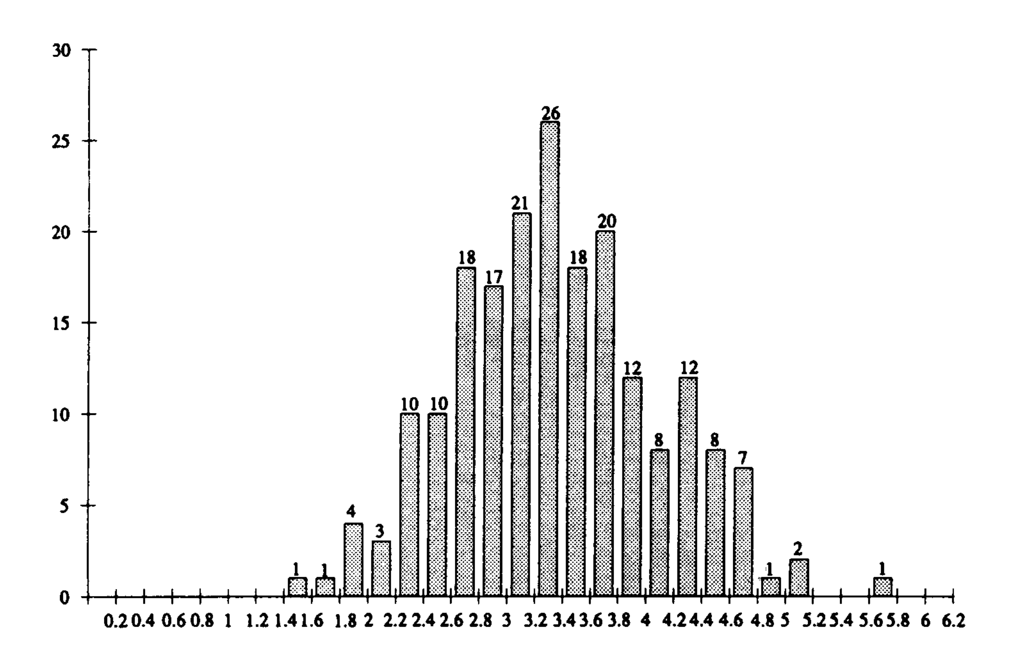
تابع جدول (٥ ـ ١)

رقم العينة	قياسات العينة	Σx_i	$\bar{x}$	رقم العينة	قياسات العينة	Σx_i	$\bar{x}$
149	1,3,6,1,3	14	2.8	175	2;4,2,2,2	12	2.4
150	6,3,1,4,6	20	3.8	176	4,6,6,6,2	24	4.8
151	3,6,6,1,3	19	3.8	177	3,6,5,4,4	22	4.4
152	3,5,2,6,2	18	3.6	178	2,3,4,4,3	16	3.2
153	3,1,2,2,5	13	2.6	179	2,6,5,3,5	21	4.2
154	4,6,4,3,3	20	4.0	180	6,3,5,2,1	17	3.4
155	1,4,2,4,3	14	2.8	181	4,3,2,2,1	12	2.4
156	5,5,4,6,4	24	4.8	182	3,5,2,2,3	15	3.0
157	4,1,4,4,3	16	3.2	183	4,3,6,1,2	16	3.2
158	3,2,1,5,5	16	3.2	184	5,5,1,6,2	19	3.8
159	5,6,1,3,5	20	4.0	185	6,2,3,3,2	16	3.2
160	2,5,6,3,3	19	3.8	186	1,4,4,4,2	15	3.0
161	1,4,2,5,3	15	3.0	187	5,6,3,6,4	24	4.8
162	4,2,4,3,5	18	3.6	188	5,1,3,5,3	17	3.4
163	1,2,5,2,6	16	3.2	189	4,4,1,3,5	17	3.4
164	1,1,3,5,2	12	2.4	190	5,3,1,2,4	15	3.0
165	3,5,3,4,5	20	4.0	191	1,1,1,6,1	10	2.0
166	3,1,2,2,4	12	2.4	192	4,5,4,4,6	23	4.6
167	2,4,3,5,2	16	3.2	193	5,2,6,6,6	25	5.0
168	2,6,3,5,3	19	3.8	194	5,6,5,5,5	26	5.2
169	5,4,3,1,1	14	2.8	195	6,5,1,6,4	22	4.4
170	6,2,6,6,6	26	5.2	196	4,2,3,4,6	21	4.2
171	1,5,5,1,1	13	2.6	197	5,2,4,2,2	15	3.0
172	3,5,5,3,1	17	3.4	198	2,3,3,3,6	18	3.6
173	1,2,2,3,1	9	1.8	199	6,1,4,5,2	18	3.6
174	2,1,4,1,2	10	2.0	200	2,3,1,1,4	11	2.2

لتوضيح الفكرة الأساسية التي تتضمنها نظرية النهاية المركزية، والتي نعرضها في العبارة المبسطة التالية :

(٥ ـ ٥ ـ ١) الفكرة الأساسية لنظرية النهاية المركزية

إذا سحبنا عينات عشوائيـة حجم كل منها n ، من مجتمع متـوسطه μ وانحرافه المعياري σ محدودان، فإن توزيع متـوسط العينة $\bar{x}$ يتطابق تقريبـا مع التوزيع الطبيعي بمتوسط يساوي μ وانحراف معياري $\frac{\sigma}{\sqrt{n}}$. وستزداد دقة التقريب كلما ازداد n .



شكل (٥ ـ ٨) مدرج التكرار لمتوسطات العينات المائتين المسحوبة من مجتمع قذف حجر النرد

ويمكن إعادة صياغة النظرية لتتفق مع $\sum_{1}^{n} X_i$ بدلا من $\bar{X}$. فنقول إن توزيع $\sum_{1}^{n} X_i$ يسعى أيضا إلى أن يصبح طبيعيا بمتوسط يساوي $n\,\mu$ وانحراف معياري $\sigma\sqrt{n}$ ، وذلك عندما يصبح n كبيرا جدا .

وتبدو أهمية نظرية النهاية المركزية من زاويتين ، فهي توضح أولا نزوع العديد من المتغيرات العشوائية إلى أن يكون توزيعها ، بصورة تقريبية ، هو التوزيع الطبيعي . إذ يمكن ، مثلا ، أن نتصور طول الإنسان حصيلة عدد كبير من المؤثرات العشوائية ، مثل طول الأب ، وطول الأم ، والمورثات (وعددها كبير) ، ونشاط الغدة أو الغدد ذات العلاقة بالطول ، والبيئة أو المحيط بأنواعه ، والتغذية ، إلخ . وإذا كانت آثار هذه العوامل ، تضاف بعضها إلى بعض ، لتنتج واقعا معينا بالنسبة إلى طول الإنسان فعندئذ

يمكن اعتبار الطول كحصيلة لعدد كبير من المتغيرات العشـوائية. وهكذا تنطبق نظرية النهاية المركزية، ويكون توزيع متغير الطول هو، على وجه التقربب، التوزيع الطبيعي، وذلك بصرف النظر عن تـوزيع أي من المتغيرات العشوائية التي تؤثـر في تحديد الطول. وهـذه بالطبع محاولـة للتعليل، ليس أكثـر، إذ أن مـا يجري في الواقـع غير معروف لنـا بصورة دقيقة، ولكن ما يمكن قوله، على كل حال، هو إن نظرية النهاية المركزية توضح سبب وجود العـديد من المتغيرات العشـوائية التي نصـادفها في حيـاتنا العـامة، والتي نعتبر أن توزيعها الاحتمالي هو التوزيع الطبيعي.

ومن زاوية أخـرى نجد أن العطاء الأكثـر أهمية لنظريـة النهاية المركـزية، يتعلق بمسألـة الاستقـراء الاحصـائي. فـالعـديـد من الإحصـاءات التي تستخـدم للقيـام باستقراءات حـول معلمات توزيع (وهي تمثل خصائص مهمة لمجتمـع القياسات) مثل p، احتمال النجـاح في التـوزيع الثنـائي، أو μ متـوسط التـوزيع الطبيعي إلخ. هـذه الإحصاءات تأخذ شكل مجموع لقياسات العينـة أو متوسط هذه القياسات. وإذا كان الحال كذلك، وكـانت n كبيرة بكفايـة، فيمكننا اعتبار التـوزيع الطبيعي تقريبـا جيدا للتوزيع الاحتمالي لـذلك الإحصاء. وهـو ما تمس الحاجة إليـه عند القيـام بأي استقراء إحصائي. وسنجد في الفقرات القادمة العديد مـن الاستخدامات المفيدة للغاية لنظرية النهاية المركزية.

والسـؤال الذي يفـرض نفسه هنـا، هو: كم يجب أن يبلغ حجم العينـة n حتى يصبح التقريب النـاشيء عن تطبيق نظرية نهايـة المركزية تقـريبا جيدا من وجهـة النظر العملية؟

ولسوء الحظ لا يوجد جواب عام ومحدد تماما لهذا السؤال. ويتعلق الأمر بالتوزيع الاحتمالي الموافق للمجتمع الذي جاءت منـه العينة، وبالغايـة من استخدام التقريب،

وهكذا. وغالبا ما يكون لكل حالة حكمها، معتمدين، بصورة رئيسة، على الخبرة السابقة والتجربة. ونشعر بكثير من الراحة عند النظر إلى مثال قذف حجر النرد المذكور أعلاه، فقد لاحظنا أن المدرج التكراري للقيم الـ 200 لـ $\bar{X}$ قريب من شكل الجرس بالرغم من أن حجم العينة الذي استخدمناه لم يتعد الخمس، وبالرغم من أن التوزيع الذي تأتي منه العينات هو خط أفقي (انظر الشكل (٣ ـ ٣)) وبعيد جدا عن شكل الجرس. وبصورة عامة، يمكن القول إنه كلما كانت درجة التناظر في التوزيع الذي نعاينه عالية كان التقريب جيدا حتى في عينات صغيرة الحجم.

تمارين (٥ ـ ٤)

١) بالإشارة إلى التمرين ١١ من مجموعة التمارين (٣ ـ ١)، لنفرض أن الشخص يقوم بـ 250 رحلة في السنة إلى عمله. وليكن $\bar{Y}$ متوسط عدد الإشارات الحمر التي يواجهها في الرحلة الواحدة، احسب $E(\bar{Y})$ ، $V(\bar{Y})$، ثم احسب $P(\bar{Y} \geq 1.5)$

٢) في مدينة معينة 1/3 الأسر ليس لديها سيارة، و 1/3 الأسر لديها سيارة واحدة، و 1/6 الأسر لديها سيارتان، و 1/12 من الأسر لديها ثلاث سيارات، و 1/12 من الأسر لديها أربع سيارات، ليكن X عدد السيارات للأسرة الواحدة:

أ ـ احسب $E(X)$ ، $V(X)$.

ب ـ احسب $E(\bar{X})$ ، $V(\bar{X})$ حيث $\bar{X}$ متوسط عينة عشوائية من 100 أسرة.

جـ ـ إذا كان لكل سيارة خمس عجلات فما المتوسط والانحراف المعياري لعدد العجلات للأسرة الواحدة.

د ـ احسب بصورة تقريبية $P(\bar{X} < 1)$.

٣) تذبح مضافة عربية كل يوم 1 ، 2 ، 3 ، أو 4 خراف باحتمالات هي، على الترتيب، 0.4 ، 0.3 ، 0.2 ، 0.1. ما هو الحد الأدنى لعدد الخراف التي ستلبي باحتمال لا يقل

عن 0.99 حاجة المضافـة من الذبائح لفترة 120 يوما؟ (نفترض أن حاجـة المضافة في يوم مستقلة عن حاجتها في يوم آخر).

٤) متوسط الوزن في قطيع ضخم من الخراف هو 8.2 كغ بتباين يساوي 4.84 كغ٢ . ما احتمال أن يقع متوسط الوزن في عينة عشوائية من 80 خروفا بين 8.3 و 8.4 كغ؟

(٥ ـ ٦) تقريب التوزيع الثنائي بالتوزيع الطبيعي

رأينا في الفصل السـابق عدة تطبيقات للتـوزيع الثنائي إقتضت جميعهـا حساب احتمال أن يأخذ X ، وهو عدد النجـاحات من بين n تكرارا، قيمـة معينة أو يقع ضمن فترة معينة، وقـد اقتصرنا هنـاك على أمثلة تكـون n صغيرة فيها، وذلك بسبـب مشقة الحسابات عندما تكون n كبيرة. ولنفرض، مثلا، أننا في حاجة لحساب احتمال وقوع X ضمن فترة معينة ، حيث $n = 1000$ ، فمع أن مثل هذا العمل ليس مستحيـلا، إلا أنه ممتنع إلى الحد الذي نريد معه تجنب الغوص في الحسابات. وتقدم نظرية النهاية المركزية حلا لهذه المشكلة. ذلك لأنه يمكن النظر إلى عدد النجاحات X كمجموع يحقق شروط نظريـة النهاية المركـزية. فإذا اصطلحنا على أن يـوافق النتيجة S (أو النجاح) العـدد 1 ويوافق النتيجة F (أو الفشل) العدد صفر. فعندئذ تكـون نتائج التكرارات المستقلة لـ n عبارة عن متتالية من المتغيرات المستقلة $X_1, X_2, ..., X_n$ ، حيث يأخذ كل X_i إما القيمة 1 أو القيمة صفر. ويكون عدد النجاحات X هو بالضبط عدد مرات ورود الـ 1 في تلك المتتالية أو مجموعها. أي أن

$$X = \sum_{i=1}^{n} X_i$$

وبما أن كل X_i يتوزع وفق التوزيع الثنائي النقطي أو توزيع بيرنوللي، [انظر مطلع الفقرة (٤ ـ ٢) ونهاية الفقرة (٤ ــ ٧)] فتصبح نتائج التكرارات المستقلة الـ n وهي

$X_1, X_2, \ldots, X_n$ عينة عشوائية من مجتمع بيرنوللي، ويصبح X مجموع هذه العينة. ووفقا لنظرية النهاية المركزية يكون التوزيع التقريبي لـ X ، في حالة n كبيرة بكفاية، هو التوزيع الطبيعي بمتوسط يساوي $n\,p$ وتباين يساوي $n\,p\,q$. وبالتالي يمكن استخدام جدول التوزيع الطبيعي المعياري لحساب احتمالات تتعلق بالمتغير X ، ولكن بصورة تقريبية.

مثال (٥ ـ ١٣)

لنأخذ التوزيع الثنائي في حالة $n = 10$ ، $p = 1/2$. وعندئذ يكون $\mu = np = 5$ و$\sigma = \sqrt{npq} = 1.58$ احسب $p(2 \le X \le 4)$ باستخدام التوزيع الثنائي أولا ثم باستخدام التوزيع الطبيعي لحساب قيمة تقريبية.

الحل

$$P(2 \le X \le 4) = \binom{10}{2}\left(\frac{1}{2}\right)^{10} + \binom{10}{3}\left(\frac{1}{2}\right)^{10} + \binom{10}{4}\left(\frac{1}{2}\right)^{10} = 0.3662$$

وهذا الاحتمال هو مجموع مساحات المستطيلات المقامة فوق 2 و 3 و 4 في المدرج الاحتمالي (انظر الشكل (٥ ـ ٩)) وإذا اعتبرنا X كأنه، على وجه التقريب، متغير $N(5, 2.5)$ ، فإن نظرة سريعة إلى الشكل (٥ ـ ٩) ستوضح أن المساحة تحت منحنى الكثافة الطبيعي من 2 إلى 4 تهمل النصف الأيسر من مساحة المستطيل المقام فوق 2 ، والنصف الأيمن من مساحة المستطيل المقام فوق 4 ، وأن التقريب سيكون أفضل لو أخذنا بدلا من $P(2 \le X \le 4)$ ، العبارة $P(2 - 1/2 \le X \le 4 + 1/2)$. ولكن:

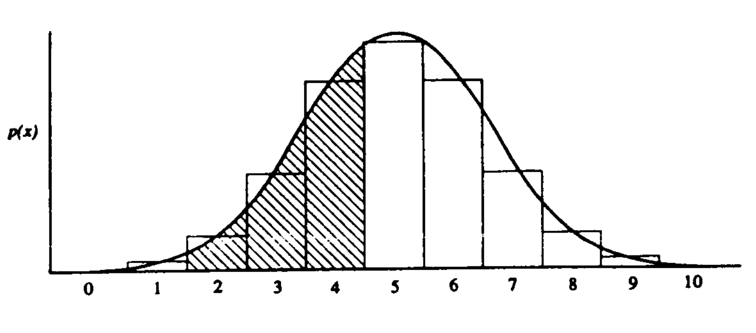
$$P(1.5 \le X \le 4.5) = F\left(\frac{4.5-5}{1.58}\right) - F\left(\frac{1.5-5}{1.58}\right)$$

$$= F(-0.316) - F(-2.215)$$

$$= 1 - F(0.316) - [1 - F(2.215)]$$

$$= F(2.215) - F(0.316) = 0.9866 - 0.6240$$

$$= 0.3626$$



شكل (٥ ـ ٩) تقريب التوزيع الثنائي بالتوزيع الطبيعي

والقيمة الناتجة صحيحة إلى رقمين عشريين بالرغم من أن n لا تتجاوز العشرة. ويعود الفضل في جودة التقريب هنا إلى تناظر التوزيع الثنائي في حالة $p = 0.5$ ، وإلى تعديل فترة تغير X ، مأخوذا كمتغير طبيعي مستمر، بحيث تغطي تماما المستطيلات الموافقة للحادثة التي نحسب احتمالها. وتسمى إضافة أو طرح 1/2 ، عملية تصحيح من أجل الاستمرار.

وعندما يكون n صغيرا و p قريبا من الصفر، أو قريبا من الواحد، فإن شكل المدرج الاحتمالي سيكون ملتويا بشدة (أي تتجمع معظم المساحة إلى جانب $X = 0$ أو إلى جانب $X = n$ ، على الترتيب) وبالتالي سيكون بعيدا جدا عن وضع التناظر. وفي مثل هذه الحالات سيكون التقريب سيئا ما لم تكن n كبيرة بكفاية.

مثال (٥ ـ ١٤)

موثوقية قطعة إلكترونية هي احتمال أن نختار واحدة من كومة إنتاج فنجدها تؤدي المهمة التي صممت من أجلها. أحسب احتمال أن نجد ما لا يقل عن 27 قطعة لا تعمل من بين عينة عشوائية تتضمن 1000 قطعة وذلك تحت الفرض بأن الموثوقية هي 0.98.

الحـــــل

المسألـة هي مسألة تـوزيع ثنائي فكل قطعـة تختارها إمـا أن تعمل أو لا تعمل. وإذا اعتبرنـا نتيجـة «القطعـة لا تعمل» نجـاحـا، يكـون $p = 0.02$ ويكون المطلـوب حساب:

$$P(X \geq 27) = \sum_{x=27}^{1000} \binom{1000}{27} (0.02)^{x} (0.98)^{1000-x}$$

والحساب الدقيق لهذه النتيجة يتطلب جهدا كبيرا. وبـاستخدام تقريب التوزيع الطبيعي نحسب المسـاحـة تحت المنحنى الطبيعي إلى اليمين من $X = 26.5$ (لاحظ أنه ينبغي استخدام $X = 26.5$ بدلا من $X = 27$ بحيث تشمل المستطيل الاحتمالي المقـام فوق النقطة $X = 27$). وذلك بـاعتبـار أن X يتبع على وجـه التقـريب، التـوزيع الطبيعي بمتوسط يساوي

$$\mu = np = 1000 \times 0.02 = 20$$

وانحراف معياري

$$\sigma = \sqrt{n\ p\ q} = \sqrt{1000 \times 0.02 \times 0.98} = 4.43$$

وهكذا نجد قيمة تقريبية للاحتمال المطلوب:

$$\begin{aligned} P(X \geq 26.5) &= P\left(\frac{X-\mu}{\sigma} \geq \frac{26.5-\mu}{\sigma}\right) \\ &= P\left(Z \geq \frac{26.5-20}{4.43}\right) = P(Z > 1.4) \\ &= 1 - F(1.4) = 1 - 0.9292 = 0.0708 \end{aligned}$$

مثال (٥ ـ ١٥)

اختبرنا لقاحـا جديدا ضد الزكام. وقـد أعطي اللقاح لمائة شخص، وروقبوا من حيث إصابتهم بالزكـام لمدة سنة. وقد نجا 68 منهم من الإصابة بالـزكام. ولنفرض أننا

نعلم من معلومات سابقة أن احتمال عدم الإصابة بالزكام هي بصورة طبيعية وبدون استخدام اللقاح 0.5 . أية نتائج يمكنك استخلاصها من هذه التجربة حول فعالية اللقاح؟

الحل

لنحسب احتمال نجاة 68أو أكثر من الإصابة بالزكام تحت الفرض بأن $p = 0.5$ ، أي أن اللقاح لم يكن له أي تأثير، فنجد باستخدام التقريب الطبيعي :

$$\mu = n\,p = 100\,(0.5) = 50\ ;\ \sigma = \sqrt{50 \times 0.5} = 5\ ,$$
$$P(X \geq 68) = P\left(Z \geq \frac{67.5 - 50}{5}\right) = 1 - F(3.5) = 0.0002$$

لقد قمنا بالحسابات مفترضين أن اللقاح غير فعال، وأن العدد 68 الذي حصلنا عليه، وهو أكبر من المتوقع تحت هذا الفرض، كان محض مصادفة. ولكن الاحتمال الناتج صغير جدا، وهو يعني ، عمليا، أنه لو كان ما افترضناه صحيحا وكررنا التجربة نفسها عددا كبيرا جدا من المرات فإننا سنجد نتيجة كالنتيجة التي حصلنا عليها، أو أفضل، في تجربتين من كل عشرة آلاف تجربة، وهذا يثير الكثير من الريبة في صحة ما افترضناه، ويدعو إلى الاعتراف بفعالية اللقاح في الوقاية من الزكام.

مثال (٥ ـ ١٦)

يتضمن امتحان خمسين سؤالا من النوع متعدد الاختيارات، ولكل سؤال ثلاثة أجوبة مقترحة، واحد منها فقط هو الجواب الصحيح؛ ولكي ينجح الطالب لا بد له من الإجابة بصورة صحيحة على عشرين سؤالا على الأقل.

ا ـ احسب احتمال نجاح طالب غير مؤهل يختار جوابه عن كل سؤال عشوائيا.

ب ـ مع بقاء عدد الأسئلة ودرجة النجاح كما هي، كم يجب أن يكون عدد

الاختيارات المطروحة أمام كل سؤال ليصبح احتمال نجاح طالب يختار جوابه عشوائيا أقل من 0.01؟

جـ ـ في حال وجود اختيارين فقط، كم يجب أن تكون درجة النجاح بحيث لا يزيد احتمال نجاح طالب يختار جوابه عشوائيا على الواحد في المائة؟

الحل

ا ـ ليكن X عدد الأجوبة الصحيحة، فلدينا $p = 1/3$ ، $n = 50$ ، والمطلوب $P(X \geq 20)$. وباستخدام التقريب الطبيعي نجد:

$$\mu = np = 50 \times \frac{1}{3} = \frac{50}{3} \;,\; \sigma = \sqrt{npq} = \sqrt{\frac{100}{9}} = \frac{10}{3}$$

$$P(X \geq 20) \approx P\left(Z \geq \frac{19.5 - \frac{50}{3}}{\frac{10}{3}} \right) = 1 - F(0.85) = 1 - 0.8023$$
$$= 0.198$$

ب ـ

$$P(X \geq 20) = P\left(Z \geq \frac{19.5 - 50p}{\sqrt{50p(1-p)}} \right) \leq 0.01$$

$$F\left(\frac{19.5 - 50p}{\sqrt{50P(1-p)}} \right) \geq 0.99$$

$$\frac{19.5 - 50p}{\sqrt{50p(1-p)}} \geq 2.33$$

والمطلوب قيمة p التي تحقق هذه المتباينة وتجعل المقدار $19.5 - 50p$ موجبا كما ينبغي أن يكون. وبتربيع الطرفين والإصلاح نجد:

$$2771.45\,p2 - 2221.45\,p + 380.25 \geq 0$$

وهذه تتحقق إذا كان $p > 0.55$ أو $p < 0.248$. ولكن قيم p الأكبر من 0.55 مرفوضة لأنها تجعل $19.5 - 50p$ سالبا. وبما أن عدد الاختيارات هو بالضرورة عدد

صحيح فعلينا أخذ أول نسبة تقل عن 0.248 ويكون جداؤها بعدد صحيح مساويا للواحد تماما. والنسبة المطلوبة هي إذا 0.2 ، وهذا يعني أن عدد الاختيارات المطروحة أمام كل سؤال ينبغي أن تكون خمسة.

جـ ـ المطلوب تحديد عدد صحيح a يحقق المتباينة التالية:

$$P(X \geq a) \leq 0.01$$

حيث $n = 50$ ، $p = 1/2$ ، وبالتالي $\mu = 25$ ، $\sigma = \sqrt{12.5} = 3.54$.

$$P(X \geq a) \approx P\left(Z \geq \frac{a - \frac{1}{2} - 25}{3.54}\right) \leq 0.01$$

أي

$$F\left(\frac{a - 25.5}{3.54}\right) \geq 0.99$$

$$\frac{a - 25.5}{3.54} \geq 2.33 \Leftrightarrow a = 34 .$$

تمارين (٥ ـ ٥)

١) عند تصالب حبتي بازيلاء لكل منها زوج من المورثات (أحمر، أبيض) يُتوقَّع أن تكون زهور ربع النسل بيضاء. إذا فحصنا 64 نبتة ناتجة عن مثل هذا التصالب فما احتمال أن نجد 16 منها بالضبط ذات زهور بيض؟

٢) نسبة القطع غير الصالحة التي تنتجها آلة هي 20% . أحسب بصورة تقريبية احتمال أن تتضمن عينة عشوائية من 400 قطعة من إنتاج هذه الآلة أكثر من 96 قطعة غير صالحة؟

٣) نقذف حجر نرد 300 مرة، ونعتبر الحصول على 1 أو 2 «نجاحا». استخدم تقريب التوزيع الثنائي بالتوزيع الطبيعي لحساب احتمال ألا يحيد عدد النجاحات عن 100 بأكثر من 15 .

٤) في بلدة كبيرة يعطي نصف الناخبين، عادة، أصواتهم للمرشح A . ويسأل كل من 20 باحثا إحصائيا عينة عشوائية من 16 من الناخبين عن المرشح المفضل. استخدم جدول التوزيع الطبيعي لحساب تقريبي لعدد الباحثين الذين تتوقع أن يفيدوا بأن أقل من 6 من عينتهم فضلوا المرشح A .

٥) يزرع رجل في حديقة منزله بذور زهور يقال أن 60% منها ينبت. إذا زرع 60 بذرة فما احتمال أن ينبت منها 15 بذرة أو أقل؟

٦) لتعيين مشرف على آلة حاسبة الكترونية تتطلب إحدى الشركات من المرشحين اجتياز اختبار كتابي. وتتألف ورقة الامتحان من 100 سؤال متعدد الاختيارات، ولكل سؤال ثلاثة أجوبة مقترحة أحدها فقط صحيح. والنجاح في الاختبار يقتضي الإجابة الصحيحة على 40 سؤالا، على الأقل. والمطلوب

ا ـ احتمال نجاح متقدم يختار الجواب على كل سؤال عشوائيا؟

ب ـ أكبر عدد من الأسئلة ينبغي أن تتضمنها ورقة الامتحان إذا أردنا لاحتمال نجاح متقدم يختار أجوبته عشوائيا أن لا يتجاوز الـ 1% ؟

٧) 25% من تلاميذ مدرسة لم يكن في سجلهم خلال عام دراسي بأكمله أي يوم غياب بسبب المرض. وفي الصف السادس من هذه المدرسة يوجد 120 تلميذا. أوجد عددا r بحيث يكون احتمال أقل من r تلميذ صف سادس بدون أي يوم غياب مرضي يساوي 0.01 أعرض الفرضيات التي اعتمدت عليها؟

٨) إذا كان 55% من الناخبين في مدينة كبيرة يؤيدون قضية فما احتمال أن تظهر عينة عشوائية من 100 ناخب من هذه المدينة أغلبية لصالح القضية؟

٩) احتمال أن نستكمل بنجاح سلسلة من العمليات في تجربة معينة هو 0.44 . إذا بدأنا 65 من مثل هـذه التجـارب بصـورة تضمن استقـلال كل تجربـة عن غيرهـا من التجارب، فما احتمال أن نستكمل بنجاح أقل من 25 منها؟ بين أنه إذا كان احتمال النجاح 0.04 فقط فإن احتمال أربع نجاحات على الأقل هو حوالي 1/4 ؟

١٠) بـالإشارة إلى التمـرين ١٥ من مجموعـة التمارين (٣ ـ ١) هل يمكنك الآن إعطـاء جواب تقريبي؟

١١) بالإشارة إلى التمرين ١٦ من مجموعـة التمارين (٣ ـ ١)، هل يمكن إعطاء جواب تقريبي؟

(٥ ـ ٧) فترة ثقة لمتوسط مجتمع طبيعي تباينه معروف

ذكرنا عبر هذا الكتاب أن الإحصاء يهدف إلى التنبؤ أو اتخاذ قرار حول خاصة من خصائص مجتمع اعتمادا على المعلومات المتيسرة من عينة نأخذها من هذا المجتمع. وكما يـوحي عنـوان الفقـرة فإن المجتمع الـذي ينبغي دراستـه هـو مجتمـع يخضع للتـوزيع الطبيعي، أو مجتمع موصوف رياضيا بنمـوذج هو النموذج الطبيعي. وأن الخاصة التي تهمنا من خصائص هذا المجتمع هي متوسطه μ ، مثلا، مع افتراض أن تباينه معروف ويساوي σ^2 . وما نريـده هنا هو تحديد فترة، أي تحديـد عددين حقيقيين، نستطيع أن نقول، بثقة عالية، إن المتوسط يقع بينهما.

لنأخـذ عينـة عشوائيـة حجمهـا n مـن هـذا المجتمـع، ولنرمـز لمقاديرهـا بـ $X_1, X_2, \ldots, X_n$ ولمتوسطها بـ $\bar{X}$. فكما رأينا في الفقرة [٥ ـ ٤ (الخاصة ٣)]، يتوزع $\bar{X}$ وفق التوزيع الطبيعي $N\left(\mu, \frac{\sigma^2}{n}\right)$. ومعـرفة التوزيع الاحتمالي لـ $\bar{X}$ يعني بـالنسبة لنا

الشيء الكثير، إذ نستطيع تقديم وصف رياضي لمجتمع القياسات الموافق لـ $\bar{X}$ ، أي للقيم كافة التي يمكن أن يأخذها المتوسط $\bar{X}$ لو أننا قمنا بأخذ عدد هائل من العينات المختلفة ذات الحجم n من هذا المجتمع. وسيسمح لنا هذا التوزيع بالإجابة بيسر وسهولة على أسئلة هامة من النوع: ما نسبة العينات التي يتجاوز متوسطها قيمة محددة؟ أو يقل عن قيمة محددة؟ أو يقع بين عددين محددين؟ الخ. وبصورة عامة، يمكننا اعتمادا على معلوماتنا من الفقرة (٦ ـ ٣) أن نكتب:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

و

$$P\left(- Z_{\alpha / 2} < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < Z_{\alpha / 2} \right) = 1 - \alpha$$

وهذا يكافىء:

$$P\left(- Z_{\alpha / 2} \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < Z_{\alpha / 2} \frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha$$

أو

$$P\left(\bar{X} - Z_{\alpha / 2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha / 2} \frac{\sigma}{n} \right) = 1 - \alpha$$

وفي هذه العبارة يمكن معرفة $Z_{\alpha/2}$ من جدول التوزيع الطبيعي المعياري حالما نحدد قيمة α ، و σ معروف، و n حجم العينة محدد سلفا. وإذا أمعنا النظر، سنجد منطوق هذه العبارة قبل أخذ العينة كالتالي:

إن نسبة $100\ (1 - \alpha)\%$ من العينات ذات الحجم n التي يمكن أخذها من هذا المجتمع ستؤدي إلى قيمة لـ $\bar{X}$ بحيث تتضمن الفترة التي تبدأ بالعدد $\bar{X} - Z_{\alpha / 2} \frac{\sigma}{\sqrt{n}}$ وتنتهي بالعدد $\bar{X} + Z_{\alpha / 2} \frac{\sigma}{\sqrt{n}}$ القيمة الصحيحة لمتوسط المجتمع μ . ويسمى $\bar{X} - Z_{\alpha / 2} \frac{\sigma}{\sqrt{n}}$ حد الثقة الأدنى و $\bar{X} + Z_{\alpha / 2} \frac{\sigma}{\sqrt{n}}$ حد الثقة الأعلى.

وتيسيرا للفهم، ولتشكيل تصور محسوس للفكرة التي نطرحها هنا، دعنا نحدد قيمة لـ α، ولتكن 0.05، وعندئذ $Z_{\alpha/2} = Z_{0.025} = 1.96$. ويكون $1 - \alpha = 0.95$. وتصبح المقولة التي تشكل منطوق العبارة الاحتمالية أعلاه كالتالي:

إن نسبة 95% من العينات ذات الحجم n التي يمكن أخذها من هذا المجتمع ستؤدي إلى قيمة لـ $\bar{X}$ بحيث تتضمن الفترة

$$\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} \; , \; \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}} \right)$$

القيمة الصحيحة لمتوسط المجتمع μ .

وعندما نأخذ العينة سنحصل على قيمة محددة $\bar{x}$ للمتغير العشوائي $\bar{X}$ ، وسنجد فترة معرفة تماما هي الفترة الممتدة بين العدد $\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}$ والعدد $\bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}$ ، وهذه الفترة إما أن تتضمن القيمة الصحيحة μ أو لا تتضمنها وليس هناك خيار ثالث. لم يعد هناك احتمالات للموقف، فقد أطلقنا طلقة على الهدف (أخذنا عينة وحددنا فترة ثقة) والنتيجة هي حتما واحدة من اثنتين فإما أننا أصبنا الهدف (الفترة تغطي μ) أو أننا لم نصبه (الفترة لا تغطي μ). وبما أننا نعلم قبل أخذ العينة أن نسبة عالية من العينات، (95% منها) تصيب الهدف، فستتولد عندنا ثقة عالية بأن العينة التي حصلنا عليها قد أصابت فعلا، مما يقترح تسمية النسبة العالية تلك «معامل ثقة». فنقول إن الفترة الناتجة هي فترة ثقة تتضمن μ بمعامل ثقة يبلغ 95%. والله سبحانه وتعالى وحده يعلم ما إذا كانت العينة التي حصلنا عليها حسنة الطالع (من بين الـ 95% التي تغطي القيمة الصحيحة للمتوسط μ) أم أنها سيئة الطالع (من بين الـ 5% التي يجانبها الصواب، إذ لا تغطي الفترة الناشئة عنها القيمة الصحيحة للمتوسط μ) .

وبالطبع يمكن أن نكون أشد تحفظا فنأخذ $\alpha = 0.01$ ويكون معامل الثقة $100\,(1 - \alpha) = 99\%$. ومن الطبيعي أن تكون الفترة التي نحصل عليها في هذه الحالة

أطول من سابقتها المقابلة لمعامل ثقة 95% . كما يمكن، على الوجه الآخر، أن نكون أقل تحفظا فنأخذ $\alpha = 0.10$ ، ويكون معامل الثقة 90% لفترة تمتاز بأنها أقصر من سابقتيها .

مثال (٥ ـ ١٧)

يمثل البيان الإحصائي التالي إنتاج عشر شجيرات من الطماطم مقاسا بالكيلوغرام .

2.3, 2.6, 2.2, 3.1, 4.0, 1.9, 2.7, 1.9, 3.3, 3.0

ونعلم أن قياسات الإنتاج في مجتمع شجيرات الطماطم يوصف بتوزيع طبيعي تباينه $\sigma^2 = 0.36$. أحسب 90% ، 95% ، و 99% فترة ثقة لمتوسط الإنتاج μ .

الحل

متوسط العينة $\bar{x}$ هو:

$$\bar{x} = \frac{2.3 + 2.6 + \ldots + 3.0}{10} = 2.7$$

90% فترة ثقة :

$$Z_{\alpha/2} = Z_{0.05} = 1.645 \text{ ، } \alpha/2 = 0.05 \text{ ، } \alpha = 0.10 \text{ ، } 1 - \alpha = 0.9$$

وتكون الفترة المطلوبة :

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.7 \pm 1.645 \frac{\sqrt{0.36}}{\sqrt{10}} = 2.7 \pm 0.31 = (2.39, 3.01)$$

95% فترة ثقة :

$$Z_{\alpha/2} = Z_{0.025} = 1.96 \text{ ، } \alpha/2 = 0.025 \text{ ، } \alpha = 0.05 \text{ ، } 1 - \alpha = 0.95$$

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.7 \pm 1.96 \frac{0.6}{3.16} = 2.7 \pm 0.37 = (2.33, 3.07)$$

99% فترة ثقة :

$Z_{\alpha/2} = Z_{0.005} = 2.58$ ، $\alpha/2 = 0.05$ ، $\alpha = 0.10$ ، $1 - \alpha = 0.99$

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.7 \pm 2.58 \frac{0.6}{3.16} = 2.7 \pm 0.49 = (2.21, 3.19)$$

لاحظ أن فترة الثقة تتسع مع ازدياد معامل الثقة .

في هذا المثال نخرج ، مثلا ، بالتقدير التالي : «بمعامل ثقة 95% يقع متوسط الإنتاج μ بين 2.33 كغ و 3.07 كغ . ولكن هب أننا اتفقنا على اعتبار منتصف الفترة قيمة تقديرية أو تقديرا لـ μ ، فهذا شيء منطقي تماما إذ نقول إن متوسط العينة $\bar{X} = 2.7$ هو تقديرنا لمتوسط المجتمع الذي جاءت منه العينة ، ولكن الاكتفاء بذلك لا يضيف أي جديد إلى ما هو معروف تاريخيا ، إذ يلجأ كل خبير يريد القيام بعملية تخمين إلى أخذ عينة تمثل المجتمع ، في رأيه ، تمثيلا جيدا ، ثم يأخذ معلومات العينة ليعممها بصورة مباشرة على المجتمع . وكأن العينة هي صورة مصغرة للمجتمع ، وليس علينا إلا تكبير هذه الصورة حتى نحصل على صورة المجتمع . ولكن ماذا عن الخطأ في هذا التقدير؟ لو رجعنا إلى العبارة الاحتمالية في مطلع الفقرة وهي :

$$P\left(- Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

لوجدنا أنها مكافئة للعبارة :

$$P\left(\left| \bar{X} - \mu \right| < Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

و $|\bar{X} - \mu|$ يمثل الخطأ المطلق للتقدير ، فهو القيمة المطلقة لحيدان التقدير عن الشيء المراد تقديره . والعبارة الإحتمالية تقول إنه باحتمال يبلغ $(1 - \alpha)$ لا يتجاوز الخطأ في هذا التقدير المقدار $\pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ ، وفي المثال السابق يمكن القول ، مثلا ، إنه في 95% من العينات الممكنة سوف لا يحيد متوسط العينة عن متوسط المجتمع بأكثر من 0.37 كغ زيادة أو نقصانا.وبعبارة أخرى ، سوف لا يتعدى الخطأ في تقديرنا إلا فيما

ندر، القيمـة 0.37 كغ زيـادة أو نقصـانـا. وسنطلق على المقـدار $Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ ، مع شيء من التجـاوز، اسم «الحد الأعلى لخطأ التقـديـر»، فهـو في حقيقة أمـره حـد أعلى تقريبي لخطأ التقدير. وسنرمز له بالرمز e ، ونكتب:

$$e = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

ولو سألنا الخبير التقليدي عن الخطأ في تقديره لعجز عن الإجابة إذ ليس لديه أية وسيلـة تسمح لـه بـذلك. وبينما يعتمد الخبير التقليـدي إعتمادا كليـا على العينـة التي أخذها فإن الإحصـائي اليوم لا يعتمد على العينة إلا كجزء من صـورة متكاملة تتضمن إلى جانب العينة المأخوذة العينات كافة التي كان يمكن الحصول عليها لو أنه كرر تجربة أخذ العينـة عددا هائلا من المرات. وهـو ما يسمى بتوزيع المعـاينة، ممثلا هنـا بتوزيع المتوسط $\bar{X}$. وهذه هي الإضافة الجديدة لعلـم الإحصاء في مسألة كهذه. (انظر الفقرة (٥ ـ ٧) والفقرة (٥ ـ ٩)).

ويتضح من عبارة e أنـه يمكننـا التحكم في حجم الخطأ من خـلال التحكم في حجم العينة n ، فالمقدار e يتناسب عكسا مع الجذر التربيعي لحجم العينـة. ولو أردنا تخفيض e إلى نصف مـا هو عليـه لاحتجنا إلى زيادة حجـم العينة إلى أربعـة أضعاف. وبـالطبع يمكننا قبل تنفيـذ البحث الإحصـائي، أي قبل أخذ العينـة، تصميم حجم العينة n بصورة تتناسب مـع مقدار الخطأ الذي يمكن التساهل فيـه. وسنوضح الفكرة بمثال.

مثال (٥ ـ ١٨)

بالإشـارة إلى المثال السـابق (٥ ـ ١٧)، لنفرض أننـا نريد تقـدير متـوسط إنتاج شجيرة الطماطم μ بحيث لا يزيد الخطأ عن 0.2 كغ إلا باحتمال زهيد لا يتجاوز الواحد في المائة. فكم يجب أن يكون حجم العينة؟

الحل

الحد الأعلى للخطأ e يساوي 0.2 ،

$$Z_{\alpha/2} = Z_{0.005} = 2.58 \qquad \alpha = 0.01$$

$$e = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 2.58 \frac{0.6}{\sqrt{n}} \leq 0.2$$

ومنه :

$$0.2\sqrt{n} \geq 1.548$$

$$\sqrt{n} \geq 7.74 \ , \ n \geq 59.9$$

أي أن حجم العينة يجب أن لا يقل عن 60 .

تمارين (٥ ـ ٦)

١) بفرض عينات عشوائية من مجتمعات طبيعية تباينها معروف ، أوجد فترات ثقة للقيمة الحقيقية μ لمتوسط المجتمع بمعامل الثقة المبين في كل حالة :

ا ـ $n = 9$ ، $\bar{X} = 4$ ، $\sigma^2 = 16$ ، معامل الثقة 90% .

ب ـ $n = 100$ ، $\bar{X} = 29$ ، $\sigma^2 = 49$ ، معامل الثقة 95% .

جـ ـ $n = 64$ ، $\bar{X} = 4$ ، $\sigma^2 = 100$ ، معامل الثقة 99% .

٢) مجتمع طبيعي انحرافه المعياري $\sigma = 0.75$. كم يجب أن يكون حجم عينة مأخوذة من هذا المجتمع كي لا يزيد الحد الأعلى للخطأ عن 0.4 ، وذلك باحتمال 0.95 ؟

٣) نعلم أن الخطأ المرتكب في قياس طول ، عند استخدام جهاز لقياس الأطوال ، يتوزع وفق التوزيع الطبيعي ، بمتوسط يساوي الصفر ، وانحراف معياري 1 مم .

ا ـ أحسب احتمال أن يقل الخطأ عند استخدام الجهاز لمرة واحدة عن 0.5 مم .

ب ـ إذا استخدم الجهاز بصورة مستقلة 9 مرات لقياس طول معين ، فاحسب احتمال أن يقع متوسط القياسات التسعة في حدود 0.5 مم من القيمة الحقيقية للطول .

٤) مجتمع طبيعي انحرافه المعياري $\sigma = 0.75$ ، كم يجب أن يكون حجم عينة مأخوذة من هذا المجتمع كي لا يزيد الحد الأعلى للخطأ عن 0.4 ، وذلك باحتمال 0.95 ؟

٥) يريدإحصائي تحديد متوسط الأجر اليومي لمستخدمي مهنة معينة. ويريد باحتمال 0.95 حداً أقصى للخطأ قدره 9 ريالات. ومن دراسات مماثلة أخرى يعلم أن بإمكانه افتراض مجتمع طبيعي انحرافه المعياري $\sqrt{650}$ ريالاً. ما هو حجم العينة التي ينبغي أن يخطط للحصول عليها؟

٦) إذا كان الانحراف المعياري لمجتمع طبيعي هو $\sigma = 5$. ما هو حجم العينة التي ينبغي أخذها حتى نطمئن باحتمال قدره 0.95 إلى عدم اختلاف متوسط العينة عن متوسط المجتمع بأكثر من الواحد؟

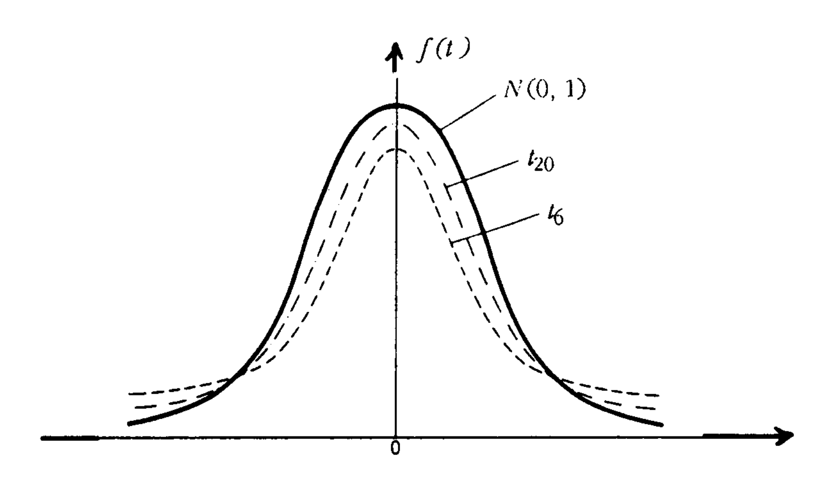
(٥ ـ ٨) فترة ثقة لمتوسط مجتمع طبيعي تباينه غير معروف وحجم العينة صغير

لو تتبعنا المناقشة في الفقرة السابقة لوجدنا أنه لا بد من تعويض σ ، التي افترضناها معروفة هناك، بتقدير لها من العينة. والتقدير الذي تمليه البداهة هو اعتماد S ، الانحراف المعياري للعينة، كتقدير لـ σ ، الانحراف المعياري للمجتمع الذي جاءت منه العينة. وهكذا يأخذ المقدار $Z = \dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}}$ ، الذي يتبع تماما التوزيع الطبيعي المعياري، الصيغة:

$$t = \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

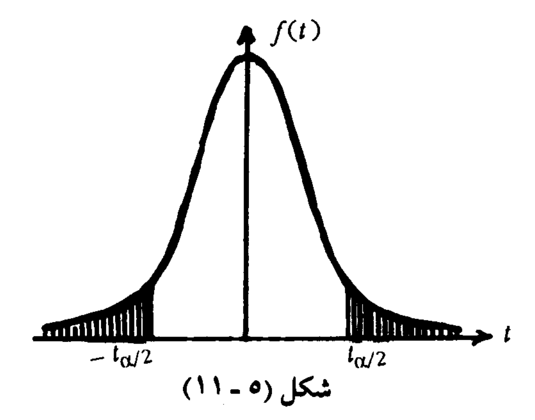
والمتغير الجديد الذي رمزنا له بـ t يحوي مركبة عشوائية في البسط هي $\bar{X}$ ومركبه عشوائية في المقام هي S. ولم يعد توزيعه هو التوزيع الطبيعي المعياري. وقد تمكن «ستيودنت» ، وهو لقب لكاتب إحصائي كان ينشر أبحاثه بتوقيع «ستيودنت»، أن

يشتق العبارة المضبوطة لتوزيع t ويسمى هذا التوزيع في كتب الإحصاء المختلفة «التوزيع t »، أو توزيع «ستيودنت». وفي الشكل (٥ ــ ١٠) نجد أمثلة من منحنيات الكثافة لهذا التوزيع. فهو متناظر حول المحور الرأسي، شأنه في ذلك شأن منحنى الكثافة الطبيعي المعياري. ويعتمد المنحنى على حجم العينة n ونصطلح على تسمية المقدار $n-1$ ، «عدد درجات الحرية» ونرمز له بـ ν (حرف يوناني ينطق نو). والجدول ٢ في الملحق يعطي القيمة الموجبة لـ t التي يقع إلى اليسار منها $(1-\alpha)$ من المساحة الكلية تحت المنحنى وذلك من أجل قيم مختلفة لـ α و ν . وسنرمز بـ $t_\alpha(n-1)$ للدلالة على قيمة المتغير t في صلب الجدول الواقعة في ملتقى السطر $n-1$ والعمود الذي عنوانه $1-\alpha$. وعلى سبيل المثال، لإيجاد $t_{0.025}(14)$ ، ندخل الجدول وفق السطر 14 ونتحرك حتى نصل إلى العمود الذي عنوانه 0.975 لنجد القيمة 2.145 .



شكل (٥ ـ ١٠) أشكال مقارنة للتوزيعين t_6، t_{20} والتوزيع $N(0, 1)$.

وإذا أخذنا عينة عشوائية حجمها n من مجتمع طبيعي $N(\mu, \sigma^2)$ حيث μ و σ^2 غير معروفين، وكالعادة رمزنا بـ $\bar{x}$ و s لمتوسط العينة وانحرافها المعياري فيمكننا إستنادا إلى تناظر التوزيع t (انظر الشكل (٥ ـ ١١)) كتابة :



شكل (٥ ـ ١١)

$$P\left(-t_{\alpha/2}(n-1) < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\right) = 1-\alpha$$

أو

$$P\left(\frac{-S}{\sqrt{n}}\, t_{\alpha/2}(n-1) < \bar{X}-\mu < \frac{S}{\sqrt{n}}\, t_{\alpha/2}(n-1)\right) = 1-\alpha$$

أو

$$P\left(\bar{X}-\frac{S}{\sqrt{n}}\, t_{\alpha/2}(n-1) < \mu < \bar{X}+\frac{S}{\sqrt{n}}\, t_{\alpha/2}(n-1)\right) = 1-\alpha$$

وتكون فترة الثقة للمتوسط μ بمعامل ثقة $100(1-\alpha)\%$ هي $\left(\bar{X}-\frac{S}{\sqrt{n}}\, t_{\alpha/2}(n-1)\ ,\ \bar{X}+\frac{S}{\sqrt{n}}\, t_{\alpha/2}(n-1)\right)$. ولوضع فترة ثقة نحسب إذا $\bar{X}$ و S من العينة وقيمة $t_{\alpha/2}(n-1)$ من جدول التوزيع t ثم نعوض.

مثال (٥ ـ ١٩)

قسنا ارتفاع خمس عشرة شجيرة باذنجان بعد فترة من زرعها فكان متوسط الارتفاع 83 سم بانحراف معياري 5.8 سم. ضع 95% فترة ثقة لمتوسط الارتفاع في المجتمع الذي اخترنا منه الشجيرات الخمس عشرة في العينة، مفترضا أن ارتفاع الشجيرة في المجتمع يخضع للتوزيع الطبيعي.

الحل

حـــدا الثقـــة همـــا $\bar{x} \pm \frac{S}{\sqrt{n}}\, t_{0.025}\,(14)$. ومـن جـدول التوزيع t نجــد $t_{0.025}\,(14) = 2.145$ ، لاحظ أنه لـو كان σ معروفا لكان هذا العدد 1.96 فقط)، وتكون فترة الثقة بمعامل 95% هي :

$$83 \pm 2.145 \times 5.8 / \sqrt{15}$$

أي من 79.8 سم إلى 86.2 سم .

وكلما ازداد حجم العينة أصبح S تقديرا أفضل لـ σ وبالتالي اقتربت قيم t من قيم المتغير الطبيعي المعياري Z الموافقة لها. ولو نظرنـا في السطور الأخيرة في جدول التوزيع t (السطور التـي تلي السطر 30) لوجـدنا أن الفـروق بين قيم t وقيم Z المقابلـة لها تصبح صغيرة، وفي السطر الأخير من الجدول حيث كتب حـذاءه الرمز "∞" تتطابق قيم t مع قيم Z المقابلة .

مثال (٥ ـ ٢٠)

وضعت عينـة من 12 فأرا تجريبيـا على نظام تغـذية معين خـلال الأشهر الثـلاثة الأولى من حياتها وقيست الزيادة في وزن كل فأر بالغرام فكانت كما يلي :

55, 62, 54, 58, 65, 64, 60, 62, 59, 67, 62, 61

والمطلوب وضع فترة ثقة بمعامل ثقة 90% لمتوسط الزيادة في الوزن خلال الأشهر الثلاثة الأولى من حياة مجتمع الفئـران الذي جاءت منـه العينة، علما أنه يمكن اعتبـار التوزيع الطبيعي توزيعا مناسبا لمتغير زيادة الوزن خلال الأشهر الثلاثة الأولى .

الحل

نحسب متوسط العينة وانحرافها المعياري فنجد $\bar{X} = 60.75$ ، $S = 3.84$ ، ولدينا $1 - \alpha = 0.90$ ، $\alpha = 0.10$ ، $\alpha/2 = 0.05$ ، ومن جدول التوزيع t نجد :

$$t_{\alpha/2}(n-1) = t_{0.05}(11) = 1.796$$

بالتعويض نجد فترة الثقة المطلوبة :

$$60.75 \pm \frac{3.84}{\sqrt{12}} \times 1.796 = 60.75 \pm 1.99$$

أي من 58.76غ إلى 62.74غ .

تمارين (٥ ـ ٧)

١) في ستة اختبارات لتجميع وتركيب قطع آلية معينة، استغرق وقت التجميع والتركيب 13 ، 14 ، 12 ، 16 ، 13 ، و 11 دقيقة. مفترضا أن زمن التجميع والتركيب يتبع التوزيع الطبيعي، ضع فترة ثقة لمتوسط الزمن الحقيقي للتجميع والتركيب بمعامل ثقة 99% .

٢) عينة عشوائية من 30 درجة من درجات اختبار للذكاء أعطي لطلاب المرحلة الثانوية، أنتجت متوسطا قدره 423 وإنحرافا معياريا $S = 68$. أوجد فترة ثقة لمتوسط المجتمع بمعامل ثقة 95% ، مفترضا أن درجات الاختبار في المجتمع تتبع التوزيع الطبيعي.

٣) وجد طبيب أسنان في فحصه الدوري لستة طلاب ابتدائي أنهم احتاجوا إلى 2 ، 3 ، 6 ، 0 ، 4 ، و 3 عمليات حشوة.

ا ـ إذا استخدم الطبيب متوسط هذه العينة كتقدير لمتوسط المجتمع الذي جاءت منه العينة فماذا يمكنه أن يقول باحتمال 0.95 عن الحد الأعلى للخطأ الذي ارتكبه؟

ب ـ ضع فترة ثقة لمتوسط المجتمع بمعامل ثقة 95% .

جـ ـ ما الفرض الذي استندت إليه في حساباتك؟

٤) أحد الثوابت الفيزيائية المهمة هو e/m نسبة شحنة الكهروب (الإلكترون) إلى كتلته.
وفي تجربة فيـزيائية لقياس هـذا الثابت أعيدت، بصـورة مستقلة، 12 مرة، كانت النتائج التالية:

$1.7604 \times 10^7 , 1.7638 \times 10^7 , 1.7609 \times 10^7$

$1.7563 \times 10^7 , 1.7556 \times 10^7 , 1.7582 \times 10^7$

$1.7526 \times 10^7 , 1.7663 \times 10^7 , 1.7624 \times 10^7$

$1.7620 \times 10^7 , 1.7605 \times 10^7 , 1.7621 \times 10^7$

ا ـ ما تقـديـرك للقيمة الحقيقة لـ e/m ؟ وما هو الحد الأعلى لخطأ هـذا التقدير باحتمال 0.95 ؟

ب ـ ضع فترة ثقة لقيمة e/m بمعامل ثقة 99% .

٥) تأتي مادة غذائية من مصنع معين في علب مكتـوب عليها «الوزن الصافي 38 أونزة» .
وقد وجد أن الوزن الذي تحتويه كل من عينة عشوائية من 6 علب كان كما يلي:

34.06, 39.65, 34.75, 40.00, 39.50, 34.25

ضع فترة ثقة لمتوسط محتوى العلبة من إنتاج المصنع من تلك المادة الغذائية، وذلك بمعامل ثقة 98% .

(٥ ـ ٩) فترة الثقة لمتوسط مجتمع في حالة عينات كبيرة الحجم

لا نفترض هنا أن المجتمع الـذي نأخذ منه العينـة مجتمع طبيعي، ولكننا نفترض أن حجم العينـة n كبير إلى الحد الذي يسمح بـالاستفادة من نظـرية النهـاية المركـزية، واعتبار توزيع $\bar{X}$ ، متوسط العينة، مطابقا تقـريبا للتوزيع الطبيعي. وبالتالي تطبيق ما جاء في الفقـرة (٥ ـ ٧) بحـذافيره. فإذا كان تبـاين المجتمع σ^2 معروفا كـانت فترة الثقة لمتوسط المجتمع μ ، بمعامل ثقة $100(1-\mu)\%$ هي

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

حيث تؤخذ قيمة $Z_{\alpha/2}$ من جدول التوزيع الطبيعي المعياري. وإذا كان التباين σ^2 غير معروف، وغالبا ما يكون الأمر كذلك في التطبيقات العملية، فإن تباين العينة S^2 يشكل تقديرا جيدا لـ σ^2، نظرا لكبر حجم العينة، مما يسمح بتعويض S بدلا من σ في فترة الثقة لتصبح:

$$\bar{X} \pm Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

وكما رأينا في الفقرة (٥ ـ ٧) نعتبر متوسط العينة $\bar{X}$ تقديرا لمتوسط المجتمع μ ويكون الحد الأعلى التقريبي لخطأ التقدير:

$$e = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

وفي حالة σ غير معروف نعوض عن σ بـ S لنجد:

$$e = Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

وبما أن معامل الثقة الأكثر استخداما في التطبيقات العملية هو المعامل 0.95 أو 95% .
فقد جرت العادة على كتابة الحد الأعلى التقريبي لخطأ التقدير على الشكل:

$$e = 1.96\ \sigma_{\bar{X}} = 1.96 \frac{\sigma}{\sqrt{n}}$$

حيث نعني بـ $\sigma_{\bar{X}}$ الانحراف المعياري لمتوسط العينة $\bar{X}$، وهو وفقا لنظرية النهاية المركزية $\frac{\sigma}{\sqrt{n}}$. وبما أن النتائج تقريبية، على أي حال، فقد جرت العادة أيضا على استخدام 2 بدلا من 1.96 ، تسهيلا للحسابات، وهكذا نكتب:

$$e = 2\sigma_{\bar{X}} = 2\frac{\sigma}{\sqrt{n}}$$

أو في حالة σ غير معروف:

$$e = 2\frac{S}{\sqrt{n}}$$

ويجدر التنويه بخطأ شائع بالنسبة إلى المبتدئين، ينبغي الانتباه إليه، وهو استخدام 2σ كحد أعلى تقريبي للخطأ بدلا من $2\sigma_{\bar{X}}$.

مثال (٥ ـ ٢١)

لنفرض أننا نرغب في تقدير متوسط الإنتاج اليومي في شركة للصناعات الكيميائية. وقد سجلنا الإنتاج اليومي لفترة $n = 60$ يوما فكان متوسط هذه العينة وانحرافها المعياري بالأطنان:

$$S = 23 \quad , \bar{X} = 941$$

والمطلوب تقدير μ متوسط الإنتاج اليومي في هذه الشركة.

الحل

التقدير الأفضل هو $\mu = 941$ طنا في اليوم. وحدود الخطأ بالزيادة أو النقصان في هذا التقدير هي:

$$\pm 2\frac{S}{\sqrt{n}} = \pm \frac{2(23)}{\sqrt{60}} = 5.94$$

(تذكر أننا عندما نستخدم العدد 2 يكون معامل الثقة 95%). وهكذا نقول، بمعامل ثقة 95%، إن التقدير 941 هو في حدود 5.94 طنا، زيادة أو نقصانا، من القيمة الحقيقية لمتوسط الإنتاج.

مثال (٥ ـ ٢٢)

نعلم أن عمر مركبة معينة من دائرة كهربائية يتبع توزيعا احتماليا ملتويا. أخذنا عينة عشوائية من 250 من هذه المركبات فكان متوسط العمر فيها 840 ساعة بانحراف معياري $S = 21.98$ ساعة. أوجد فترة ثقة تقريبية لمتوسط العمر في مجتمع إنتاج هذه المركبات، مستخدما معامل ثقة 95%.

الحل

فترة الثقة المطلوبة هي:

$$\bar{X} \pm 2\frac{S}{\sqrt{n}}$$

أي

$$840 \pm 2\,\frac{21.98}{\sqrt{250}} = 840 \pm 2.78$$

وهكذا نقدر بمعامل ثقة 95% أن متوسط العمر في مجتمع إنتاج هذه المركبات واقع بين 837.2 و 842.8 ساعة .

مثال (٥ ـ ٢٣)

نريد تقدير μ متوسط الطول في إنتاج مصنع للبراغي في حدود خطأ لا يزيد عن 1/2 مم إلا باحتمال لا يتجاوز الخمسة في المائة . فكم يجب أن يكون حجم العينة علما بأننا نعرف من سجلات الإنتاج السابقة أن الانحراف المعياري للطول يساوي 1.2 مم؟

الحـــل

$$e = Z_{\alpha/2}\,\frac{\sigma}{\sqrt{n}} = 1.96\,\frac{1.2}{\sqrt{n}} \le \frac{1}{2}$$

ومنه

$$\sqrt{n} \ge 2\,(1.96)(1.2) = 4.704$$

$$n \ge 22.1$$

أي أن حجم العينة يجب ألا يقل عن 23 .

تمارين (٥ ـ ٨)

١) من المعروف أن عمر أحد عناصر دائرة كهربائية يتبع توزيعا احتماليا ملتويا . وقد وجدنا أن متوسط العمر في عينة عشوائية من 300 عنصرا يساوي 920 ساعة بتباين يساوي 483 ساعة^2 . ضع فترة ثقة تقريبية لمتوسط العمر الحقيقي لهذا العنصر مستخدما معامل ثقة 95% .

٢) في عينة عشوائية من 100 كيس تفاح كتب عليه «الوزن الصافي 1 كغ» وجدنا أن متوسط الوزن 1002 غراما بتباين يساوي 144 غ^2 . ضع فترة ثقة تقريبية لمتوسط وزن التفاح الحقيقي ضمن الكيس الواحد وذلك بمعامل ثقة 90% .

٣) عينة عشوائية من 60 مخزنا أظهرت أن متوسط سعر الحليب $\bar{X} = 77.3$ سنتا للكيلوغرام، بانحراف معياري 4.2 سنتا. أوجد فترة ثقة لمتوسط سعر الحليب بمعامل ثقة 95% .

٤) مالك سيارة يريد أن يعرف المتوسط الأسبوعي للمسافة التي يسيرها مقاسة بالميل. وقد سجل المسافات التي قطعها في 52 أسبوعا متتاليا ووجد متوسطها 176 ميلا في الأسبوع، بانحراف معياري 96 ميلا. ضع فترة ثقة لمتوسط ما يقطعه في الأسبوع بمعامل 95% .

٥) يرغب مستشفى في تقدير عدد الأيام التي يحتاجها علاج مرضى يقع سنهم بين 25 و 34 سنة. وقد وجدت إدارة المستشفى أن متوسط عدد أيام الإقامة لعينة عشوائية من 500 مريض من هذه الفترة من العمر، يساوي 5.4 يوما بانحراف معياري 1.3 يوما. ضع فترة ثقة لمتوسط الإقامة في المستشفى لمجتمع المرضى الذي جاءت منه العينة، وذلك بمعامل ثقة 99% .

٦) لنفرض أن تباين مجتمع $\sigma^2 = 100$ ، وبمعامل ثقة 95% نريد أن يكون تقديرنا $\bar{x}$ في حدود 2.5 وحدة قياس من μ المتوسط الحقيقي للمجتمع. كم يجب أن يكون حجم العينة؟

٧) فيما يلي جدول التوزيع التكراري للعمر عند الزواج، لأقرب سنة، لـ 175 رجلا:

مركز الفئة	17.5	22.5	27.5	32.5	37.5	42.5	47.5	52.5	57.5	62.5
التكرار	28	28	43	18	9	4	2	1	0	2

ا ـ أحسب متوسط العمر عند الزواج وانحرافه المعياري .

ب ـ إذا افترضنا أن هذه الأعمار عينة عشوائية من مجتمع كبير. فاحسب بمعامل ثقة 95% فترة ثقة لمتوسط العمر عند الزواج في ذلك المجتمع .

٨) في تجربة لبنج جديد، أعطي لمائة فأر وقيس زمن الانتعاش لكل منها إلى أقرب عشر من الدقيقة، فكانت النتائج كما يلي :

45.0	58.2	55.1	52.5	61.7	52.9	70.4	62.5	71.3	50.1
84.9	60.9	35.4	64.3	75.7	48.5	41.3	53.8	66.8	37.4
32.4	50.7	82.3	71.8	66.4	49.7	51.7	56.0	88.8	64.7
77.9	41.4	52.7	53.4	57.9	51.7	55.6	44.1	85.4	67.3
87.3	52.5	46.7	48.3	60.1	66.0	77.3	46.5	54.3	52.6
53.1	67.9	55.9	64.2	68.0	48.2	41.2	56.3	79.4	80.9
58.7	49.0	51.2	70.2	54.0	74.6	51.9	42.6	95.4	51.9
83.5	70.4	76.7	47.0	55.9	43.8	49.1	60.0	38.3	44.3
63.5	45.4	57.3	54.5	73.9	64.1	80.6	68.8	73.5	84.0
65.9	58.3	59.6	59.1	46.7	51.3	44.5	54.2	63.8	56.9

والمطلوب تقدير متوسط زمن الانتعاش في المجتمع الذي جاءت منه العينة وإعطاء حد أعلى لخطأ التقدير باحتمال 0.99 .

كم فأرا تحتاج كي يكون خطأ التقدير 1 تقريبا؟

(٥ ـ ١٠) فترة الثقة لنسبة

لنفرض أن صنفا معينا A يوجد في مجتمع كبير بنسبة تساوي π . إذا سحبنا عينة عشوائية حجمها n من هذا المجتمع، وعرفنا النجاح بأنه الحصول على عنصر من الصنف A ، فإن احتمال النجاح عند كل سحب هو، عمليا، π . ونسبة النجاح في

العينة هي عدد عناصر الصنف A ولنرمز لها بـ X (أي عدد النجاحـات) مقسوما على حجم العينة n . وإذا رمزنا لنسبة الصنف A في العينة بـ p فإن :

$$p = \frac{X}{n}$$

وعندما نـاقشنا تقريب التوزيع الثنائي بـالتوزيع الطبيعي في الفقرة (٥ ـ ٦) وجـدنا أنه يمكن اعتبـار عدد النجـاحات X مجموع عينـة ، وبالتـالي تكون النسبـة p هي متـوسط العينة . وكما أن تطبيق نظرية النهاية المركزية على X يسمح لنا بالقول إن X يتوزع تقريبا وفق التوزيع الطبيعي $N(n\pi , n\pi(1-\pi))$ ، فإن تطبيق نظريـة النهايـة المركزيـة على المتـوسط p يسمح لنا بالقول إن للنسبة p (نسبة النجاح في العينة) توزيعـا مطابقا تقريبا للتوزيع الطبيعي $N\left(\pi , \frac{\pi(1-\pi)}{n}\right)$ ، حيث :

$$E(p) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{n\pi}{n} = \pi$$

$$V(p) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{n\pi(1-\pi)}{n^2} = \frac{\pi(1-\pi)}{n}$$

وذلك كله شريطة أن يكون n كبيرا (مثلا أكبر من 30 في حالة قيمة لـ π ليست قريبة من الصفر أو قـريبة من الواحـد) . وهذه النتيجـة تسمح لنا بوضع فترة ثقـة للنسبة π على الشكل التالي بمعامل ثقة $100(1-\alpha)\%$:

$$p \pm Z_{\alpha/2}\, \sigma_p = p \pm Z_{\alpha/2} \sqrt{\frac{\pi(1-\pi)}{n}}$$

حيث σ_p تعني الانحراف المعياري للنسبة p . ووجود π المجهولة في هـذه العبارة يمنع من تطبيقها . ويمكننا الاستعاضة عن π ، نسبة النجاح في المجتمع ، بتقدير لها هو p ، نسبة النجاح في العينة . (تماما كما عوضنا عن σ بـ S في الفقرة السـابقة) . وتصبح فترة الثقة لـ π كما يلي :

$$p \pm Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

ومن أجل معـامل ثقة 95% يمكن اعتبـار $Z_{\alpha/2}$ مساويـا لـ 2 بـدلا من 1.96 ، تسهيـلا للحساب . وهكذا نكتب فترة الثقة لـ π ، بمعامل ثقة 95% كما يلي :

$$p-2\sqrt{\frac{p(1-p)}{n}} \leq \pi \leq p+2\sqrt{\frac{p(1-p)}{n}}$$

مثال (٥ ـ ٢٤)

مـن بين 300 أسـرة اخترنـاها من بلدة كبيرة وجدنا 123 أسرة تمتلك تلفازا ملونا . ضـع فترة ثقـة لنسبـة الأسر التي تمتلك تلفـازا ملونـا في مجمل البلـدة . وذلك بمعامل ثقة 95% .

الحل

لدينا $n = 300$ ، $p= \frac{123}{300} = 0.41$

$$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.41 \times 0.59}{300}} = \sqrt{0.00080633} = 0.0284$$

وهكذا تكون فترة الثقة المطلوبة كما يلي :

$$0.41 \pm 2\,(0.0284) = 0.41 \pm 0.057$$

أي أن π واقعة بين 0.353 و 0.467 أو أن ما بين 35.3% إلى 46.7% من الأسر في هذه البلدة تمتلك تلفازا ملونا .

مثال (٥ ـ ٢٥)

تضمنت عينة من 250 من طلبة الجامعة 30 طالبا أعسر (يستخدم اليد اليسرى) . أعط فترة ثقة تقريبية بمعامل ثقة 95% لنسبة الطلاب العسر في الجامعة؟

الحل

$n = 250$ ، $p= \frac{30}{250} = 0.12$

$$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.12 \times 0.88}{250}} = \sqrt{0.0004224} = 0.021$$

وتكون فترة الثقة المطلوبة :

$$0.12 \pm 2\ (0.021) = 0.12 \pm 0.042$$

أي أن ما بين 7.8% إلى 16.2% من طلبة الجامعة يستخدمون اليد اليسري .

تمارين (٥ ـ ٩)

١) اختبرنا نوعا جديدا من مصابيح آلات التصوير لتقدير p ، احتمال أن ينتج المصباح الإضاءة اللازمة وفي الوقت المناسب . ووجدنا من بين 1000 مصباح أن 920 قد عملت وفقا للمواصفات المطلوبة . والمطلوب :

ا ـ تقدير p ووضع حد لخطأ التقدير (باحتمال 0.95) .

ب ـ وضع فترة ثقة للقيمة الحقيقية p بمعامل ثقة 99% .

٢) اختبرنا عينة عشوائية من 400 صماما خاصا بأجهزة الراديو، فوجدنا من بينها 40 صماما عاطلا عن العمل . ضع فترة ثقة للنسبة الحقيقية للصمامات العاطلة في مجتمع الصمامات المنتجة من هذا النوع ، مستخدما معامل ثقة 90% .

٣) حضر كيميائي مبيدا يهدف عند تطبيقه على نوع معين من الحشرات ، إلى قتل 60% منها فكم يجب أن يكون حجم العينة المستخدمة ، إذا رغب في أن يطمئن باحتمال 0.95 إلى أنه في حدود 0.02 من النسبة الحقيقية التي يهدف إليها من الحشرات الهالكة؟

٤) كم ناخبا يجب أن تضم عينة جمعناها لتقدير نسبة الناخبين الذين يفضلون مرشحا معينا ، إذا رغبنا في أن يكون التقدير صحيحا في حدود 0.005 ؟ ولنفرض أن النسبة الحقيقية ينبغي أن تقع في جوار الـ 0.5 .

الملاحق

الملحق الأول

مراجعة في بعض المعلومات الرياضية المفيدة

١ - حول خاصية التجانس في عملية الجمع

لكي نستطيع جمع كميتين أو مقدارين لابـد أن تكـون وحـدة القيـاس نفسهـا للمقـدارين وجمع 10سم و 5م غير ممكـن. ولـو ادعينـا جـدلاً أن المجمـوع الناتج 15 فلا يمكـن القول إن ال 15 هـذه هي 15 سم ولا 15م. إذا 15 مـاذا؟ في الواقع لا معنى للعـدد 15 في هـذه الحالـة. ولكن لو اتخذنـا وحدة قيـاس مشتركـة، وقلنا إن 500 سـم = 5م، أو 10 سـم=0.1م، فـالجمـع يصبح ممكنـا، والجواب هـو 500سم+10سم=510سم، أو 5م +0.1م = 5.1م. ولو سألت شخصا، ما مجموع 5 كتب و 7 أقلام؟ فأجاب 12، فإنه سيشعر بالعجز والخطأ عندما تطلب منه تحديد 12 ماذا؟ إذ لا يستطيع أن يقول 12 كتابا ولا 12 قلما. وسيعود يدرك أن 5 كتب و 7 أقلام هي 5 كتب و7 أقلام، ولا يمكن التعبير عنهما في رقم واحد.

وهذه الحقيقـة البسيطة هي خـاصة أسـاسية في الجمع بصرف النظـر عن طبيعة عمليـة الجمع. ففي الجبر لا نجمع إلا الحدود المتشـابهة. و $5x^2y$ مضافـا إليها $7x^2y$ يسـاوي $12x^2y$. تماما كأننا نقول 5 كتب و7 كتب تصبح 12 كتابا. ولكن $5x^2y$ مضافا إليها $7xy$ هي $5x^2y$ و $7xy$، تماما كأننـا نقول 5 كتب و 7 أقلام، ولا يمكـن جمعهما في حد واحـد، لأنهما غير متشـابهين. وكـذلك الأمـر، يمكـن جمع $F(0.5)$ و $3F(0.5)$

لنجد $4F(0.5)$، حيث $F(0.5)$ تعني قيمة دالة F عند النقطة 0.5. ويمكن جمع $3\sqrt{2}$ و $5\sqrt{2}$ لنجد $8\sqrt{2}$ أما $F(0.5) + 3F(1)$، أو $3\sqrt{2} + 5\sqrt{3}$، فستبقى كل منهما كما وردت دون تغيير.

وتبقى الفكرة نفسها في الكسور العادية (الأعداد النسبية). فمقام الكسر يعني أننا قسمنا الواحد الصحيح إلى عدد من الأجزاء المتساوية يساوي المقام. والبسط يعني أننا أخذنا من هذه الأجزاء عددا يساوي البسط. و 3/4 تعني ثلاثة أرباع. أي قسمنا الواحد الصحيح إلى أربعة أجزاء متساوية وأخذنا منها ثلاثة أجزاء. ومقلوب مقام الكسر يُعتبر بمثابة وحدة قياس أو شيء، والبسط يمثل عدد ما لدينا من هذه الوحدة أو هذا الشيء. و 9/5 تعني تسع مرات المقدار 1/5 أو تسع أخماس وهكذا. . .
ولجمع كسرين عاديين لابد إذا من توحيد المقامات حتى يصبح الجمع ممكنا ولجمع 3/4 و 9/5 نحول الكسرين بحيث يكون لهما المقام نفسه. ومن خواص الكسر أو العدد النسبي نعرف أن قيمته لا تتغير إذا ضربنا البسط والمقام بالعدد نفسه، وقيمة 3/4 هي نفس قيمة 15/20، وقيمة 9/5 هي نفس قيمة 36/20. إذا

$$\frac{3}{4} + \frac{9}{5} = \frac{15}{20} + \frac{36}{20} = \frac{15 + 36}{20} = \frac{51}{20} .$$

لأن 15/20 هي 15 مرة الـ 1/20، و 36/20 هي 36 مرة الـ 1/20. ومجموعهما هو 15 + 36 = 51 مرة الـ 1/20، أو 51/20.

ولا توجد مشكلة في ضرب كسرين عاديين فالجواب هو كسر بسطه جداء البسطين ومقامه جداء المقامين.

$$\frac{3}{4} \times \frac{9}{5} = \frac{27}{20}$$

وبصورة عامة:

$$\frac{a}{b} \times \frac{c}{d} = \frac{a \times c}{b \times d}$$

وينطبق على الطرح ما ينطبق على الجمع. وطرح عدد b من عدد a، أي $a - b$، هو جمع العدد السالب $(-b)$ إلى العدد a، أي $a + (-b)$. (تذكر أننا نقرأ في اللغات

الأجنبيـة من اليسار إلى اليمين). وكـذلك ينطبق على التقسيم مـا ينطبق على الضرب. وقسمة a على b ما هي إلا ضرب a بمقلوب b، $a \div b = \frac{a}{b} = a \times \frac{1}{b}$ ولتقسيـم كسر عادي على كسر عادي آخر نضرب الأول في مقلوب الثاني. وهكذا يكون:

$$\frac{a}{b} \div \frac{c}{d} = \frac{a}{b} \times \frac{d}{c} = \frac{a \times d}{b \times c}$$

وعلى سبيل المثال:

$$\frac{3}{4} \div \frac{9}{5} = \frac{3}{4} \times \frac{5}{9} = \frac{15}{36}$$

ولحساب كسر من عدد معين نضرب هذا الكسر بالعدد. ولحساب ثلث الستة، مثلا، نضرب 6 بـ 1/3 فنجد

$$\frac{1}{3} \times 6 = \frac{1}{3} \times \frac{6}{1} = \frac{6}{3} = 2$$

ولو أردت حساب خمسة أسباع الثلاثين تكتب ببساطة:

$$\frac{5}{7} \times 30 = \frac{5}{7} \times \frac{30}{1} = \frac{150}{7}$$

٢ ـ النسب المئوية

يتألف العـدد العشري من جـزئين، أحـدهما صحيح يقـع على اليسـار من الفاصلـة، والآخـر عشري يقع على اليمين من الفاصلـة. وعلى سبيل المثال، 921.534 فيه جـزء صحيح هو 921 وجزء عشري هـو 0.534. (تذكـر في الكتب الانجليزيـة وما شابهها أن النقطـة مكتوبـة بين رقمين تعني فاصلـة عشريـة، وأنها بين رمزيـن تعني إشـارة ضرب، و $a \, . \, b$ تعني $a \times b$.) ولا تتغير قيمـة العدد إذا أضفنـا مزيـدا من الأصفار على اليمين من الجزء العشري، أو على اليسار من الجزء الصحيح، 921.534 = 0921.534000 وكما أن للجـزء الصحيح منـازل (أو مـراتب) هي منـزلة الآحـاد ومنـزلـة العشرات ومنزلـة المئات . . . الخ فكـذلك للجزء العشري منـازل تبدأ بمنـزلة الجزء من عشرة، تليها منـزلة الجزء من مـائة، تليها منـزلة الجزء من ألف، وهكـذا. ونلاحظ أن الرقم يختلف مدلولـه من منزلة إلى أخرى. فالخمسة نكتبهـا في منزلة الآحاد تعني خمس

مرات الواحد الصحيح، وهي في منزلة العشرات تعني خمس عشرات أي 50، وهي في منزلة المئات تعني خمس مئات أي 500. وكذلك في الجزء العشري، نجد أن للرقم نفسه مدلول يختلف باختلاف المنزلة التي يشغلها. فالخمسة بعد الفاصلة مباشرة، أي في منزلة الجزء من عشرة، تعني خمسة أجزاء من عشرة أي 5/10 ، وهي في منزلة الجزء من مائة تعني خمسة أجزاء من مائة أي 5/100 ، وفي منزلة الجزء من ألف تعني خمسة أجزاء من ألف أي 5/1000 ، وهكذا. واصطلاح 921.534 يعني $9 \times 10^2 + 2 \times 10 + 1 \times 10^0 + 5 \times 10^{-1} + 3 \times 10^{-2} + 4 \times 10^{-3}$ أي

$$9 \times 100 + 2 \times 10 + 1 \times 1 + 5 \times \frac{1}{10} + 3 \times \frac{1}{100} + 4 \times \frac{1}{1000}$$

ولو أردنا التعبير عن هذا المجموع كلاميا لقلنا:

واحد وعشرتان وتسع مئات بالإضافة إلى خمسة أعشار وثلاثة أجزاء من مائة وأربعة أجزاء من ألف. وبالطبع سيكون من الأيسر بكثير أن نصطلح على القول تسعمائة وواحد وعشرون فاصلة خمسمائة وأربع وثلاثون، ونكتب 921.534.

ونلاحظ أيضا أن كل منزلة في العدد العشري هي عشرة أمثال تلك التي تقع إلى اليمين منها مباشرة. فالمائة هي عشر عشرات، والعشرة عشر وحدات، والواحد عشرة أجزاء من عشرة، والجزء من عشرة هو عشرة أجزاء من مائة، وهكذا. ولذلك يطلق على هذا النظام في الترقيم اسم «النظام العشري». وهذا يوضح أن ضرب عدد على شكل كسر عشري بمضاعفات العشرة، أو قسمته على مضاعفات العشرة، لا يحتاج إلا إلى إزاحة الفاصلة في اتجاه اليمين عند الضرب، أو في اتجاه اليسار عند القسمة، عددا من المنازل يساوي عدد الأصفار في مضاعف العشرة. وعلى سبيل المثال لضرب 20.604 بعشرة نأخذ الفاصلة منزلة واحدة إلى اليمين فنجد 206.04؛ ولقسمته على عشرة نأخذ الفاصلة منزلة واحدة إلى اليسار فنجد 2.0604. وفي الحالة الأولى أصبح كل جزء من عشرة واحدا $(1/10 \times 10 = 1)$ أي أن منزلة الجزء من عشرة أصبحت هي منزلة الآحاد، ولذلك أزحنا الفاصلة منزلة واحدة إلى اليمين . وفي الحالة الثانية أصبح كل واحد صحيح جزءا من عشرة $(1/10 = 10 \div 1)$ ، أي أن منزلة الآحاد أصبحت منزلة الجزء من عشرة ، ولذلك أزحنا الفاصلة منزلة واحدة إلى اليسار. وكقاعدة عامة، عند الضرب بـ

10^n نأخذ الفاصلة إلى اليمين n منزلة، وعند القسمة على 10^n (أي الضرب بـ10^{-n}) نأخذ الفاصلة إلى اليسار n منزلة.

والآن كيف نعبر عن عدد عشري في شكل نسبة مئوية؟

النسبة المئوية تعني النسبة إلى مائة، أي عدد نسبي مقامه يساوي 100. وخمسون في المائة تعني 50/100، وخمس وستون بالمائة تعني 65/100، ونصف بالمائة تعني 0.5/100 أي 5/1000. وللتعبير عن عدد عشري في صيغة نسبة مئوية نضرب العدد العشري بمائة فنحصل على العدد المطلوب، ونضيف إلى قراءته عندئذ كلمتي «في المائة».

مثال (١)

عبر عن الأعداد التالية في شكل نسبة مئوية:

0.75، 0.3، 1.21، 0.05، 0.052، .0003، 21.3، 0.3254.

الحـــل

الأجوبة المطلوبة هي على الترتيب: 75 في المائة؛ 30 بالمائة؛ 121 في المائة؛ 5 في المائة؛ 5.2 في المائة (ونقرؤها خمس واثنان من عشرة في المائة)؛ 0.03 في المائة (ونقرؤها ثلاثة من مائة في المائة)؛ 2130 في المائة، 32.54 في المائة (ونقرؤها اثنان وثلاثون وأربع وخمسون من المائة في المائة).

ولجمع الأعداد العشرية نرتب المنازل المتشابهة تحت بعضها تماما. وبالتالي تقع الفواصل تحت بعضها البعض تماما. ثم نجمع جمعا عاديا ونضع الفاصلة عندما نصل إلى موقع الفاصلة.

٣ ـ التناسب

إذا كانت المقادير a، b، c، d بحيث يكون

$$\frac{a}{b} = \frac{c}{d}$$

قلنا إنها متناسبة. وتسمى العلاقة بينها تناسبا طرفاه a و d ووسطاه c و b. ومن أهم خواص التناسب:

أ ـ $a \times d = b \times c$ ، (جداء الطرفين = جداء الوسطين).

ب ـ $\frac{a}{a \pm b} = \frac{c}{d \pm c}$

جـ ـ $\frac{a \pm b}{b} = \frac{c \pm d}{d}$

د ـ $\frac{a}{b} = \frac{c}{d} = \frac{a+c}{b+d}$

مثال (٢)

الأعداد 2، 3، 4، 6 متناسبة. اكتب التناسب وتحقق من الخواص المذكورة أعلاه.

الحـــل

الأول والرابع هما الطرفان والثاني والثالث (الواردين في الوسط) هما الوسطان.

$$\frac{2}{3} = \frac{4}{6}$$

ومن الواضح أن

أ ـ $2 \times 6 = 3 \times 4$

ب ـ $\frac{2}{2+3} = \frac{4}{4+6}$

$$\frac{2}{5} = \frac{4}{10}$$

و $\frac{2}{3-2} = \frac{4}{6-4}$

$$\frac{2}{1} = \frac{4}{2}$$

جـ ـ $\frac{2+3}{3} = \frac{4+6}{6}$

$$\frac{5}{3} = \frac{10}{6}$$

و $$\frac{2-3}{3} = \frac{4-6}{6}$$

$$-\frac{1}{3} = -\frac{2}{6}$$

د- $$\frac{2}{3} = \frac{4}{6} = \frac{2+4}{3+6} = \frac{6}{9}$$

مثال (٣)

صندوق يتضمن 5 كرات سود، و 6 كرات بيض. ما نسبة الكرات السود إلى الكرات البيض؟ وما نسبة الكرات البيض في الصندوق؟

الحـــل

نسبة الكرات السود إلى الكرات البيض هي 5/6.

نسبة الكرات البيض في الصندوق هي نسبة عدد الكرات البيض إلى مجموع عدد الكرات في الصندوق أي

$$\frac{5}{5+6} = \frac{5}{11}$$

مثال (٤)

في صندوق 15 كرة بعضها أبيض والآخر أسود. إذا كانت الكرات تتوزع بين اللونين الأبيض والأسود بنسبة 2:1 فاحسب عدد الكرات من كل لون.

الحـــل

مجموع الحصص 1+2=3

ويكون 1/3 الكرات أبيض و 2/3 الكرات أسود.

عدد الكرات البيض = $15 \times 1/3 = 5$

عدد الكرات السود = $15 \times 2/3 = 10$

مثال (٥)

قسمنا ستة آلاف ريال بين ثلاثة أشخاص بنسبة 1:3:2 فما هي حصة كل منهم؟

الحــــل

مجموع الحصص 6=2+3+1

مقدار الحصة الواحدة 1000 = 6000/6

وتكون حصة الأول= 2 × 2000 = 2000 ريال

حصة الثاني = 3 × 1000 = 3000 ريال

حصة الثالث = 1 × 1000 = 1000 ريال .

أو نقول إن حصة الأول تشكل 2/6 من المبلغ وحصة الثاني 3/6 من المبلغ وحصة الثالث 1/6 من المبلغ. أي أن:

حصة الأول = 2/6 × 6000 = 2000 ريال

حصة الثاني = 3/6 × 6000 = 3000 ريال

حصة الثالث = 1/6 × 6000 = 1000 ريال

٤ ـ العمليات الأساسية في المجموعات وقانونا دي مورغان

الاحتـــــــواء

نقول إن المجموعة A محتواة في المجموعة B، ونكتب $A \subset B$ ، إذا كـان كل عنصر ينتمي إلى A ينتمي أيضا إلى B.

وفي حـال وجـود عنصر واحـد على الأقل ينتمي إلى B ولا ينتمي إلى A نسمي A مجموعة جزئية فعلية من B. ومن الواضح أن كل مجمـوعة محتواة في نفسها، ونـرمز لهذه الحقيقة بكتابة $A \subseteq A$.

المجموعتان المتساويتان

نقول إن المجموعتين A و B متساويتان إذا كانت $A \subset B$ و $B \subset A$. الشرط الأول $A \subset B$ يقتضي أن كـل عنصر ينتمـي إلى A ينتمي أيضـا إلى B ، والشرط الثاني $B \subset A$ يقتضي أن كل عنصر ينتمي إلى B ينتمي أيضا إلى A. وهذا يعني بوضوح تطابق عناصر المجموعتين .

اتحاد مجموعتين

اتحاد مجموعتين A ، B هو مجمـوعة تتضمن كافة العنـاصر التي تنتمي إلى واحدة منهما على الأقل . ونرمز له بـ $A \cup B$.

تقاطع مجموعتين

تقاطع مجموعتين A ، B هو مجموعة تتضمن كافة العناصر التي تنتمي إليهما معا . ونرمز له بـ $A \cap B$.

ومن الواضح أنه يمكن التعبير عن انتماء عنصر إلى $A \cup B$ بقولنا إنه ينتمي إلى A أو B. ويمكن التعبير عن انتماء عنصر إلى $A \cap B$ بقولنا إنه ينتمي إلى A و B وإذا لم يكن هناك أي عنصر مشترك بين A و B كان تقاطعهما خاليا . ونرمـز للمجموعة الخالية بـ ϕ. ومن الواضح أن المجمـوعة الخالية ϕ محتواة في أي مجموعـة أخرى ($\phi \subset A$ حيث A أي مجموعة غير خالية) .

المجموعتان المنفصلتان

إذا كـان تقاطع المجمـوعتين A ، B خاليـا ، أي $A \cap B = \phi$، قلنـا إن المجمـوعتين منفصلتان .

الفرق بين مجموعتين

الفرق بين مجموعتين A ، B، ونرمز له بـ $A\text{-}B$ (أو A/B) هو مجموعة تتضمن كافة العناصر التي تنتمي إلى A ولا تنتمي إلى B.

مكملة مجموعة

مكملة مجموعة A هي مجموعـة تتضمن كافة عنـاصر المجموعة الشـاملة التي لا تنتمي إلى A. ونرمز لها بـ $\bar{A}$ (أو A^c).

ومن الــواضح أن $\bar{A}$ تشكل نفي A. ولذلك نقرؤها أحيانا «ليس A» . ومكملة مجموعـة A ليست إلا الفـرق بين المجمـوعة الشاملـة ، ولنرمز لهـا بـS، وبين A، أي أن $A = S - A$. وهذا بالاضـافة إلى نص التعريف (٢ـ ٥ ـ ٧) يـوضحان أنه يمكن التعبير عن الفرق بين مجموعتين A ، B كتقاطع بين A ومكملة B. وهكذا نكتب :

$$A - B = A \cap \bar{B}$$

قانونا دي مورغان

ومن الخواص المهمة لعمليتي الاتحاد والتقـاطع أن كلا منهما تتـوزع على الأخرى بمعنى أن :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

و

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

وأن

$$\overline{(A \cup B)} = \bar{A} \cap \bar{B}$$

و

$$\overline{(A \cap B)} = \bar{A} \cup \bar{B}$$

وتسمى العلاقتان الأخيرتان «قانوني دي مورغان» . وتقول الأولى منهما إن مكملة اتحاد مجموعتين هي تقـاطع مكملتيهما . وتقول الثانيـة إن مكملة تقـاطع مجموعتين هي اتحاد مكملتيهما . وتسهيلا لحفظ هاتين العلاقتين نلاحظ أنه للانتقال من الطرف الأيسر إلى الطرف الأيمن في كل منهما نتبع القاعدة التالية :

نقلب اتجاه الرمز الفاصل بين المجموعتين (إذا كانت الفتحة إلى الأعلى تصبح إلى الأسفل والعكس) ثم نستعيض عن كل مجموعة بمكملتها .

مثال (٦)

لتكن المجموعة الشاملة مجموعة الحروف في الكلمات أبجد، هوز، حطي.

ولنرمز لها بـ S. ولتكن $A =$ {أ، ب، هـ، د، ط}

$B =$ { أ، هـ، ز، جـ، ط، ي}

$S =$ {أ، ب، جـ، د، هـ، و، ز، ح، ط، ي}

i) هل A محتواة في B؟

ii) أكتب $A \cup B$، $A \cap B$، $A - B$، $B - A$، $\bar{A}$، $\bar{B}$.

iii) اكتب $A \cap \bar{B}$ وقارنها مع $A - B$، واكتب $B \cap \bar{A}$ وقارنها مع $B - A$.

iv) لنعرف المجموعة $C =$ {أ، و}. تحقق من قانوني التوزيع.

v) اكتب $\overline{A \cup B}$، $\bar{A} \cap \bar{B}$، $\overline{A \cap B}$، $\bar{A} \cup \bar{B}$، ثم تحقق من صحة قانوني دي مورغان.

الحـــل

i) A غير محتواة في B لوجود عنصر د ينتمي إلى A ولكنه لا ينتمي إلى B.

$A \ni$ د ولكن $B \not\ni$ د

ii) $A \cup B =$ {أ، ب، هـ، د، ط، ز، جـ، ي}

لكتابة اتحاد A، B نكتب عناصر A ثم نضيف إليها عناصر B ما لم تكن ذُكرت سابقا، لأنه عند التعبير عن مجموعة، لا نكرر ذكر أي عنصر من عناصرها.

$A \cap B =$ {أ، هـ، ط}

ولكتابة تقاطع A، B نستعرض عناصر A واحدا فآخر ونضع في $A \cap B$ ما نجده منها واردا في B.

$A - B =$ {ب، د}

وللحصول على $A - B$ نلغي من عناصر A كل ما كان منها مشتركا مع B. وبصورة مماثلة نجد:

$B - A =$ {ز، جـ، ي}

$\bar{A} = S - A =$ { أ، ب، جـ، د، هـ، و، ز، ح، ط، ي } - { أ، ب، هـ، د، ط }

= { جـ، و، ز، ح، ي }

ولكتابة $\bar{A}$ نلغي عناصر A من المجموعة الشاملة ونأخذ كل ما تبقى منها.

وبصورة مماثلة نجد:

$\bar{B} =$ { ب، د، و، ح }

(iii

$A \cap \bar{B} =$ { ب، د } $= A - B$

$B \cap \bar{A} =$ { ز، جـ، ي } $= B - A$

iv) نحسب الطرف الأيمن فنجد

$B \cup C =$ { أ، هـ، ز، جـ، ط، ي، و }

$A \cap (B \cup C) =$ { أ، ب، هـ، د، ط } $\cap$ { أ، هـ، ز، جـ، ط، ي، و }

= { أ، هـ، ط }

نحسب الطرف الأيسر فنجد

$A \cap C =$ { أ }

$(A \cap B) \cup (A \cap C) =$ { أ، هـ، ط }

وهو يساوي الطرف الأيمن.

v) بالعودة إلى النتائج في ب نجد:

$\overline{A \cup B} =$ { أ، ب، هـ، د، ط، ز، جـ، ي }^{c} = { و، ح }

$\bar{A} \cap \bar{B} =$ { جـ، و، ز، ح، ي } $\cap$ { ب، د، و، ح } = { و، ح } $= (A \cup B)^c$

مما يحقق قانون دي مورغان الأول.

ولدينا أيضا:

$\overline{A \cap B} =$ { أ، هـ، ط }^{c} = { ب، جـ، د، و، ز، ح، ي }

و

$\bar{A} \cup \bar{B} =$ { جـ، و، ز، ح، ي } $\cup$ { ب، د، و، ح }

$\bar{A} \cup \bar{B} =$ { ب، د، و، ح، جـ، ز، ي } $= \overline{A \cap B}$

مما يحقق قانون دي مورغان الثاني.

حاصل الضرب الديكارتي

الحاصل الديكارتي لمجموعتين A ، B هو مجموعـة كل الأزواج المرتبة التي يمكن تشكيلها بأخذ العنصر الأول من A والعنصر الثاني من B.ونرمز له عادة بـ$A\times B$.

مثال (٧)

لدينا $A = \{a,b,c\}$ ، $B=\{1,2,3\}$.اكتب الحاصل الديكارتي $A\times B$.

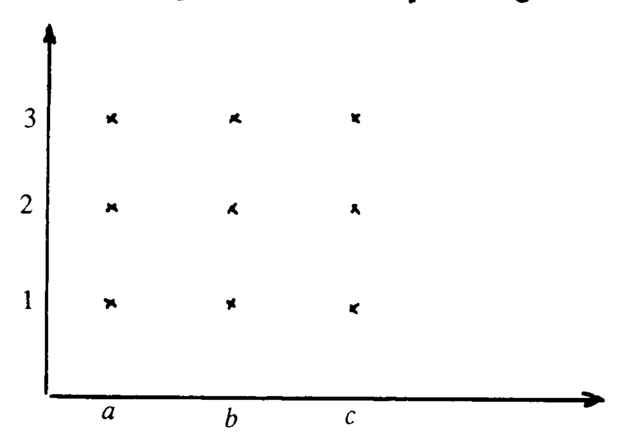
الحـــل

$$A\times B = \{(a,1),(a,2),(a,3),(b,1),(b,2),(b,3),(c,1),(c,2),(c,3)\}$$

ونلاحظ أنه إذا كـان عدد عناصر A مساويا n وعد د عنـاصر B مساويا m فإن عدد عناصر الحاصـل الديكارتي $A\times B$ يساوي $n\times m$. ويمكن تمثيل كل زوج مـرتب بنقطـة في المستوى الاحـداثي ، حيث يشكل العنصر الأول من الـزوج المرتب الاحـداثي السيني للنقطـة ويمثل العنصر الثاني الاحداثـي الصادي لها . ونحصل عنـدئذ على ما يسمى «بيان الحاصل الديكارتي» .

مثال (٨)

ارسم بيان الحاصل الديكارتي للمجموعتين A ، B في المثال (٧) .



شكل (١): بيان الحاصل الديكارتي $A\times B$.

٥ ـ التطبيق والصورة العكسية

تعريف التطبيق

التطبيق f المعرف على مجموعة A إلى مجموعة B هو توافق يقابل بموجبه كل عنصر من A عنصر واحد وواحد فقط من B . ونكتب رمزيا

$$A \xrightarrow{f} B$$

وتسمى المجموعة A مجال التطبيق (أو ساحة التطبيق) وتسمى المجموعة B المجال المصاحب. ونكتب $f(a) = b$ للدلالة على أن العنصر a من المجال A يوافقه العنصر b من المجال المصاحب B . ونقول إن b هو صورة a وفق التطبيق f .

مثال (٩)

في المثال (٧) عرف تطبيق f_1، f_2 على A إلى B .

الحل

أي قاعدة توافق يقترن بموجبها كل عنصر من A بعنصر واحد من B تشكل تطبيقا. وعلى سبيل المثال يمكننا تعريف f_1 كما يلي:

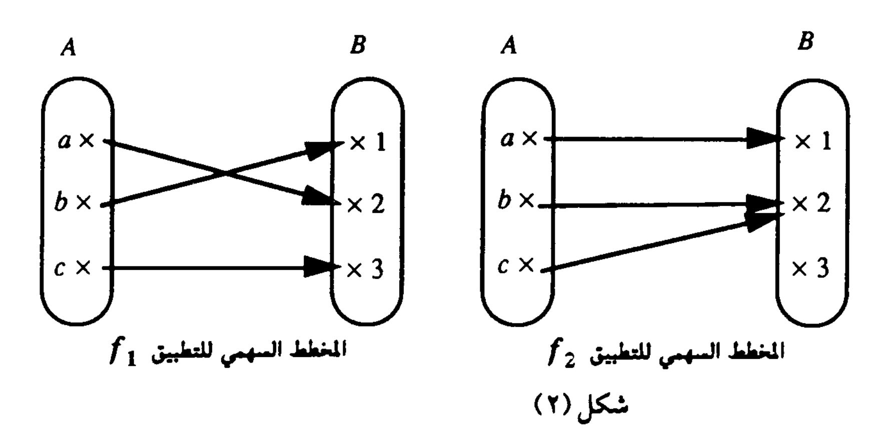
$$f_1 = \{(a,2), (b,1), (c,3)\}$$

وتقول قاعدة التوافق أو التطبيق الذي رمزنا له بـ f_1 إن العنصر a من A يقابله أو يوافقه العنصر 2 من B؛ والعنصر b من A يوافقه العنصر 1 من B؛ وأخيرا يوافق العنصر c من A العنصر 3 من B . أو أن صورة a وفق f_1 هي 2 وصورة b هي 1 وصورة c هي 3. ويمكن تعريف تطبيق آخر f_2 كما يلي:

$$f_2 = \{(a,1), (b,2), (c,2)\}$$

ونلاحظ أن شرطي التعريف محققان. فلكل من عناصر A الثلاثة عنصر مقابل واحد من B. هنا a يقابله 1؛ و b يقابله 2؛ و c يقابله 2 أيضا. أي أن 1 هي صورة لـ a و 2 صورة لكل من b و c . أما 3 فليست صورة لأي عنصر من A .

ويمكن تـوضيح التطبيق بمخطط سهمي ينطلق فيه من كل عنصر من المجـال سهم واحد ينتهي بـالعنصر المقابل (بصورته) من المجـال المصاحب. وفي الشكل (٢) نجد المخطط السهمي لكل من f_1 وf_2 في المثال (٥).



المخطط السهمي للتطبيق f_1 المخطط السهمي للتطبيق f_2

شكل (٢)

وتجدر ملاحظة أن التطبيق هو مجموعة من الأزواج المرتبة التي لا يتكرر فيها أبدا العنصر الأول. وتشكل العناصر الأولى مجتمعـة مجموعة المجال دون زيـادة أو نقصان. وعلى المخطط السهمي نقـول إنه ينطلق من كل عنصر من عنـاصر المجال سهم واحـد وواحد فقط.

تعريف الصورة العكسية

الصورة العكسية لعنصر d، مثلا، من المجال المصاحـب، هي مجموعة عناصر المجـال التي صورتها وفق f هي d. (أي مجموعة عنـاصر المجال التي انطلق منها سهم إلى d) ونرمز لها بـ $f^{-1}(d)$.

مثال (١٠)

في المثال السابق اكتب الصورة العكسية لكل عنصر من عناصر B وفـق f_1^{-1} ووفق f_2^{-1}.

الحـــل

$$f_1^{-1}(3) = c \ , \ f_1^{-1}(2) = a \ , \ f_1^{-1}(1) = b$$

ونلاحظ أن f_1^{-1} يمثل بدوره تطبيقا من B إلى A **يسمى التطبيق المعاكس. ولا يصح هذا إلا عندما يكون f تقابلا.** أي الحالة التي يكون فيها كل عنصر من B صورة لعنصر واحد وواحد فقط من A وبصورة مماثلة نجد أن

$$f_2^{-1}(2) = \{b, \ c\} \ , \ f_2^{-1}(1) = a$$

ونلاحظ أن الصورة العكسية للعنصر 2 من B هو مجموعة مؤلفة من عنصرين b، c من عناصر A ذلك لأن 2 هي صورة لكل من b و c وفق f_2. أما f_2^{-1} (3) فهي خالية ونكتب f_2^{-1} (3) $= \phi$ ، ولا يمثل f_2^{-1} تطبيقا لأنه لا يحقق شرطي التعريف، إذ يقابل العنصر 2 من B عنصران من A هي b و c، وكذلك لا صورة للعنصر 3 من B .

تعريف الدالة العددية

إذا عُرِّف تطبيق f من مجموعة جزئية من R، مجموعة الأعداد الحقيقية، إلى مجموعة جزئية أخرى منها، فإننا نسمي مثل هذا التطبيق دالة عددية ذات متغير حقيقي. وكثيرا ما نهمل عند تعريف دالة عددية، المجال والمجال المصاحب، ونعطي فقط قاعدة الربط بشكل علاقة رياضية، $y = f(x)$، ونعتبر في هذه الحالة أن مجال الدالة هو أوسع مجموعة جزئية من R يمكننا أن نُجري عليها العمليات الداخلة في القاعدة f. ويسمى المجال في هذه الحالة مجموعة التعريف أو ساحة التعريف ويسمى المجال المصاحب مدى الدالة.

٦ ـ رمز المجموع Σ وخواصه

تستخدم العلوم الرياضية ومختلف العلوم التجريبية الرموز للدلالة على مقادير أو مسميات وأشياء من طبائع مختلفة. فمثلا نرمز لمقدار أو لقياس عددي بـ x، ونرمز لمجموعة بـ A . ونرمز لعنصر من مجموعة بـ a، الخ. وفي بحث أو دراسة معينة ينبغي أن نستخدم رموزا مختلفة للدلالة على قياسات أو أشياء مختلفة، وذلك تجنبا للالتباس. وفي دراسة فيزيائية، مثلا، لو رمزنا لشدة التيار بـ x، فيجب المحافظة على هذا الرمز في الدراسة بأكملها. وحيثما وردت x ضمن هذه الدراسة فإنها تعني شدة التيار. وقد

نحتاج في دراسة واحدة إلى عدد هائل من الرموز، وربما ما لا نهاية له من الرموز، ولا يمكن لحروف أبجدية أو حروف مختلف الأبجديات المعروفة في العالم أن تفي بالحاجة. ولذلك نلجأ إلى استخدام الحرف نفسه x، مثلا، عددا هائلا من المرات دون أن نقع في التباس، وذلك بإضافة دليل رقمي تحت الحرف، فنكتب x_1 و x_2، مثلا، كرمزين مختلفين. ومع أننا استخدمنا هنا الحرف الأبجدي نفسه، إلا أننا ميزنا بين x وآخر بالدليل 1 ملحقا بالأول وبدليل آخر ملحقا بالثاني ونقرؤه x واحد، x اثنان، الخ. ويفتح لنا استخدام الحرف مع دليل رقمي ملحق به آفاقا واسعة، بحيث يمكن استخدام الحرف x نفسه عددا لانهائيا من المرات، لنكتب، مثلا، المتوالية اللانهائية من المقادير:

$$x_1, x_2, x_3, \ldots, x_n \quad , \ldots$$

يسمى x_n الحد العام للمتوالية. وعندما يتغير n متخذا عناصر مجموعة الأعداد الطبيعية $\{1,2,3,\ldots\}$، كقيم له نحصل على متوالية لانهائية من المقادير لكل منها رمز مختلف. ولم نستخدم فيها إلا حرفا واحدا من حروف الأبجدية هو x. ويمكن أن نكتب متوالية أخرى

$$y_1, y_2, y_3, \ldots, y_n \quad , \ldots$$

ومتوالية أخرى

$$z_1, z_2, \ldots, z_n \ , \ldots$$

وهكذا.

لنفرض الآن أننا نريد التعبير عن مجموع ستة مقادير هي x_1، x_2، x_3، x_4، x_5، x_6. فمن الطبيعي كتابة:

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6$$

وسنتفق الآن على التعبير عن هذا المجموع بصورة مختصرة فنكتب:

$$\sum_{i=1}^{6} x_i$$

ونقرؤها «مجموع x_i من $i = 1$ إلى $i = 6$». ونستخدم هنا الحرف اليوناني الكبير $\sum$، ويسمى «سيجما»، ليدل على كلمة «مجموع». والدليل i يسمى «متغير الجمع» وهو

يتغير هنا من $i = 1$ إلى $i = 6$. ويسمى x_i «الحد العام». وللحصـول على الحد الأول في المجموع نضع $i = 1$ في الحد العام، وفي الحد الثـاني نضع $i = 2$. وهكذا. وتفصل بين الحدود المختلفة إشارة + بالطبع مادمنا نعبر عن مجموع عدد من المقادير.

مثال (١١)

أ ـ اكتب بالتفصيل ما تعنيه الرموز:

$$\sum_{t=1}^{3} t x_t^t \quad , \quad \sum_{j=1}^{4} j(j-1) \quad , \quad \sum_{i=1}^{3} i(y_i - 1) \quad , \quad \sum_{i=1}^{5} x^i$$

ب ـ استخدم إشارة المجموع $\sum$ للتعبير عن كل من المجاميع التالية :

$$x_1 + x_2 + x_3 + x_4$$

$$y_1 + y_2 + y_3 + y_4 + y_5$$

$$x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4 + x_5 y_5$$

الحـــل

أ ـ

$$\sum_{i=1}^{5} x^i = x + x^2 + x^3 + x^4 + x^5 \ ,$$

$$\sum_{i=1}^{3} i(y_i - 1) = (y_1 - 1) + 2(y_2 - 1) + 3(y_3 - 1)$$

$$\sum_{j=1}^{4} j(j-1) = 1(1-1) + 2(2-1) + 3(3-1) + 4(4-1) \ ,$$

$$\sum_{t=1}^{3} t x_t^t = x_1 + 2x_2^2 + 3x_3^3$$

ب ـ

$$\sum_{i=1}^{5} x_i y_i \quad , \quad \sum_{i=1}^{5} y^i \quad , \quad \sum_{i=1}^{4} x_i$$

وتنبغي ملاحظـة أن الشيء الوحيد من عبارة الحد العام الـذي يتغير من حد إلى آخـر من حدود المجمـوع هـو متغير الجمع i. وبهذا المعنى يمكن اعتبـار أي كميـة لا

تتضمن متغير الجمع في حكم الثابتة. ويمكن أن يرد متغير الجمع دليلا أو معاملا أو قوة أو مقدارا قائما بذاته الخ.

خواص رمز المجموع Σ

الخاصة الأولى

$$\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i$$

وهو واضح من كون:

$$\sum_{i=1}^{n} cx_i = cx_1 + cx_2 + \ldots + cx_n$$
$$= c(x_1 + x_2 + \ldots + x_n)$$
$$= c\sum_{i=1}^{n} x_i$$

الخاصة الثانية

$$\sum_{i=1}^{n}(x_i + y_i + z_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} z_i$$

وهو واضح أيضا من الخاصتين التبديلية والتجميعية لعملية جمع الأعداد الحقيقية. فلدينا

$$\sum_{i=1}^{n}(x_i + y_i + z_i) = (x_1 + y_1 + z_1) + (x_2 + y_2 + z_2) + \ldots + (x_n + y_n + z_n)$$
$$= (x_1 + x_2 + \ldots + x_n) + (y_1 + y_2 + \ldots + y_n) + (z_1 + z_2 + \ldots + z_n)$$
$$= \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} z_i$$

الخاصة الثالثة

$$\sum_{i=1}^{n} c = nc$$

وهذا واضح من حقيقة أن الحد العام c لا يتضمن متغير الجمع i، فهو ثابت من حد إلى آخر. أي أن:

$$\sum_{i=1}^{n} c = \underbrace{c + c + \ldots + c}_{n \text{ مرة}} = nc$$

مثال (١٢)

تطبيقا لخواص الرمز Σ يمكن كتابة ما يلي :

$$\sum_{i=1}^{n} (x_i + c)^2 = \sum_{i=1}^{n} (x_i^2 + 2cx_i + c^2)$$

$$= \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} 2cx_i + + \sum_{i=1}^{n} c^2$$

$$= \sum_{i=1}^{n} x_i^2 + 2c \sum_{i=1}^{n} x_i + nc^2 .$$

ونجد أيضا :

$$\sum_{i=1}^{n} (x_i - 3y_i)^2 = \sum_{i=1}^{n} (x_i^2 - 6x_i y_i + 9y_i^2)$$

$$= \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} 6x_i y_i + 9 \sum_{i=1}^{n} y_i^2$$

$$= \sum_{i=1}^{n} x_i^2 - 6 \sum_{i=1}^{n} x_i y_i + 9 \sum_{i=1}^{n} y_i^2$$

٧ ـ محور الأعداد الحقيقية ـ الانسحاب وتغيير سلّم القياس

يمكن تمثيل الأعداد كنقاط على مستقيم موجه نسميه محورا . ولهذا الغرض نرسم مستقيما كما في الشكل (٣)، نتخذ عليه اتجاها موجبا إلى اليمين، ثم نحدد عليه نقطة تدعى عادة مبدأ الفصول، وتقابل العدد «صفر» وتقع الأعداد الموجبة إلى اليمين من مبدأ الفصول والأعداد السالبة إلى اليسار منه . ومع تبني طول معين ليمثل وحدة قياس (السنتمتر، مثلا، كما في الشكل (٣)) تصبح كل نقطة من المحور ممثلة لعدد حقيقي واحد هو عدد وحدات القياس التي تفصل بينها وبين مبدأ الفصول مسبوقا بإشارة موجبة إذا كانت النقطة إلى اليمين من مبدأ الفصول . أما إذا وقعت النقطة إلى اليسار

من مبدأ الفصول فـالعدد الحقيقي الذي تمثله هـو عدد وحدات القياس الفـاصلة بينها وبين المبدأ مسبوقا بإشـارة سالبة. وبالعكس كل عدد حقيقي تقـابله نقطة واحدة على هذا المحور. وبذلك نستكمل تمثيل مجموعـة الأعداد الحقيقية هندسيا، ويسمى المحور المدرج النـاتج محور الأعـداد الحقيقية. ونـلاحظ أن تحديـد المحـور، في الشكل (٣)، كمحـور للأعـداد الحقيقية قـد اُستكمل بعد تحديـد اتجاه عليه واتخاذ نقطـة من نقاطـه كمبدأ للفصول وتبني السنتمتر كوحدة للقياس.

−6 −5 −4 −3 −2 −1 صفر +1 +2 +3 +4 +5 +6 +7

شكل (٣): محور الأعداد الحقيقية

ولو أننـا أضفنا إلى كل عـدد المقدار 2 فإن النقطـة التي تمثل العدد «صفـر»، أي تشكل مبدأ الفصول في الشكل (٣)، ستصبح ممثلة للعدد 2، والنقطة التي تمثل العدد 2- ستصبح ممثلة للعدد «صفر»، أي مبدأ الفصـول الجديد، والنقطة التي تمثل العدد 4 ستصبح ممثلـة للعدد 6، وهكذا... ويبـدو بوضـوح أن هذا التغيير في تمثيـل النقاط للأعـداد يكافىء تماما انسحابـا إلى اليمين بمقدار وحدتي قيـاس. فكأن النقطة 2 - قد زحفت لتحتل موقع نقطـة الأصل. وفي المقابل، لو طرحنا مـن كل عدد المقدار 2 (أي أضفنا إلى كل عدد 2-) فإن النقطة التي كـانت تمثل العدد 2 ستصبح الآن ممثلة للعدد «صفر» (نقطة الأصل)، والنقطة التي كـانت تمثل الصفر (نقطة الأصل) ستصبح ممثلة للعدد 2-، وقس على ذلك. وهـذه التغييرات في تمثيل النقـاط لـلأعداد تكـافىء تمامـا انسحابا إلى اليسار بمقدار 2.

وبصورة عامة، نقول إن إضافة أو طرح عدد ثابت b، مثلا، إلى كل عدد يكافىء انسحاب التدريج بكامله مسافة b وحدة قياس إلى اليمين، إذا كان b موجبا، ومسافة b - وحدة قياس إلى اليسار إذا كان b سالبا. وإن إضافة أو طرح عدد ثابت تبدو وكأنها تغيير في اختيار نقطة الأصل.

لنفرض الآن أن لدينا مجمـوعة من القياسات {2, 3, 4, 5} فهي تحتل مـواقع معينة على محـور الأعـداد في الشكـل (٣)، ولو أضفنـا إلى كـل منهـا العدد 4 فإنها ستصبح

{6, 7, 8, 9} وتحتل مواقع جديدة في الشكل (٣) هي المواقع الناتجة عن انسحاب المواقع الأساسية بمقدار 4 وحدات قياس إلى اليمين. وإضافة العدد 4 - (أي طرح العدد 4) إلى كل منها يؤدي إلى انسحابها يسارا بمقدار 4. ونلاحظ أن المواقع النسبية للقياسات الأربعة من بعضها البعض لم تتغير بعد عملية الانسحاب.

لنضرب الآن كل عدد بالمقدار 2، مثلا، فالنقطة التي تمثل العدد صفر ستبقى في مكانها بدون تغيير، ولكن النقطة التي كانت تمثل العدد 2 ستصبح الآن ممثلة للعدد 4، والنقطة التي تمثل 4 - ستصبح الآن ممثلة للعدد 8 -، وهكذا. . . وهذه التغييرات في تمثيل النقاط للأعداد هي بالضبط ما سنحصل عليه لو أننا غيرنا وحدة القياس من السنتمتر إلى نصف السنتمتر (أي ضربنا وحدة القياس بـ 1/2). إذ لو اتخذنا نصف النستمتر وحدة لقياس المسافة في الشكل (٣) لكانت النقطة التي تمثل العدد 1 حاليا ممثلة للعدد 2، ولكانت النقطة الممثلة للعدد 3 - حاليا ممثلة للعدد 6 -، وهكذا. . . وفي المقابل لو أننا قسمنا كل عدد على 2 (أي ضربنا كل عدد بـ 1/2) فإن النقطة التي كانت تمثل العدد 2 ستصبح الآن ممثلة للعدد 1، والنقطة التي كانت تمثل العدد 4 - ستصبح الآن ممثلة للعدد 2 -، وهكذا. . . وهذه التغييرات في تمثيل النقاط للأعداد تكافىء تماما ما سنحصل عليه لو أننا غيرنا وحدة القياس إلى 2 سنتمتر بدلا من السنتمتر الواحد (أي ضربنا وحدة القياس بـ 2). وفي الحالتين نقول إننا غيرنا سلّم القياس.

وبصورة عامة، نقول إن ضرب كل عدد بمقدار ثابت موجب a يكافىء تغيير سلّم القياس بضرب وحدة القياس بـ $1/a$ (تصغيراً لها إذا كان a أكبر من الواحد وتكبيراً لها إذا كان a أصغر من الواحد). وتسمى عملية الضرب بعدد موجب عملية تغيير في سلّم القياس.

لنعد إلى مجموعة القياسات {2, 3, 4, 5} فإذا ضربنا كلا منها بـ 2 فإنها ستحتل مواقع جديدة هي النقاط المقابلة لـ {4, 6, 8, 10} وتجدر ملاحظة أن المواقع النسبية للقياسات الأربعة بعضها من بعض قد تغيرت الآن. فتغيير سلم القياس يغير من المواقع النسبية لجملة من القياسات بعضها من بعض، ولكن عملية الانسحاب لا تؤثر في تلك المواقع النسبية.

ولو افترضنا الآن أن الرمز x يمثل عددا دارجا على محور الأعداد فإن إجراء التحويل من x إلى y وفق العلاقة:

$$y = x + b$$

هو تعبير جبري عن عملية انسحاب بمقدار b . وإجراء التحويل من x إلى y وفق العلاقة:

$$y = ax$$

هو تعبير جبري عن عملية تغيير في سلم القياس. ومن الواضح الآن أن إجراء تحويل من x إلى y وفق العلاقة:

$$y = ax + b$$

تعني القيام بعملية تغيير في سلم القياس (الضرب بمقدار a) ثم القيام بعملية انسحاب للنقاط الناتجة بمقدار b.

مثال (١٣)

لدينا الأعداد

9200, 8200, 7200, 6200, 5200, 4200, 3200, 2200, 1200

إذا حولنا وفقا للعلاقة:

$$y = \frac{x - 5100}{1000} = \frac{1}{1000}\ x - \frac{5200}{1000} = 0.001\ x - 5.2$$

فإن الأعداد المعطاة تتغير، على الترتيب، إلى:

4, 3, 2, 1, 0, – 1, – 2, – 3, – 4

مثال (١٤)

لدينا الأعداد

0.011, 0.012, 0.013, 0.014, 0.015, 0.016, 0.017

إذا حولنا وفقا للعلاقة:

$$y = 1000x - 14$$

فإن الأعداد المعطاة تتغير، على الترتيب، إلى:

– 3, – 2, – 1, 0, 1, 2, 3

٨ ـ أنواع القياسات

يتعامل الإنسان مع ثلاثة أنواع من المتغيرات. وسنصطلح، بصورة عامة، على تسمية القيم التي يفترضها متغير «قياسا». ومجموعة من القياسات هي، على وجه العموم، مجموعة من القيم لمتغير أو أكثر.

والنوع الأول من المتغيرات هو المتغير الوصفي، وهو متغير يصنف جملة من العناصر وفق صفات محددة، كأن نصنف السكان في مدينة الرياض وفقا للصفات التالية:

سعودي، عربي غير سعودي، غير عربي

والمتغير هنا هو متغير الجنسية وهذه الصفات الثلاث هي قيمه الممكنة، إذ يأخذ بالنسبة لكل فرد يسكن الرياض واحدة وواحدة فقط من هذه القيم الثلاث (هنا الصفات الثلاث). وإذا رمزنا لهذا المتغير بالرمز x، واقتفينا جنسية شخص يسكن الرياض فوجدناه «غير عربي» قلنا إن المتغير x أخذ عند هذا الشخص القيمة «غير عربي». وإذا سألنا شخصا ثانيا وثالثا ووجدناهما سعوديين نقول إن قيمة المتغير x لكل منهما هي «سعودي» وهكذا. وإحدى الصفات المميزة للمتغير الوصفي هي أن مجموعة القيم التي يأخذها تصنف جملة من العناصر أو الأشياء إلى أصناف بحيث ينتمي كل عنصر (أو شيء) منها إلى صنف واحد وواحد فقط. أو بعبارة أخرى لابد أن ينتمي كل عنصر (أو شيء) منها إلى صنف من تلك الأصناف ولا يمكنه أن ينتمي إلى صنفين أو أكثر في آن واحد. وهكذا تجزىء قيم المتغير الوصفي جملة من العناصر (أو الأشياء) إلى أجزاء منفصلة بعضها عن بعض، انفصالا تاما. وفي المثال السابق لايمكن وجود مقيم في الرياض يتصف بأنه «سعودي» و«غير عربي». أو أنه «عربي غير سعودي» و«غير عربي» الخ.

والنوع الثاني من المتغيرات هو متغير ترتيبي. والمتغير الترتيبي يتميز بكل ما يتميز به المتغير الوصفي بالاضافة إلى توفر نوع من الترتيب الذي يمكن إضفاءه على

الصفات التي تشكل قيم المتغير الممكنة. فلو فرضنا، مثلا، أن متغيرا y يمثل التقدير الذي ناله طالب من طلاب فصل معين. فالقيم الممكنة لهذا المتغير هي ممتاز، جيد جدا مرتفع، جيد جدا، جيد مرتفع، جيد، مقبول مرتفع، مقبول، ضعيف. وبما أن هذه الصفات مرتبطة بمقياس كمي هو الدرجة العددية التي نالها الطالب فإن هذا يمنح ترتيبا لهذه الصفات من الأعلى إلى الأدنى أو العكس، فنقول إن أعلى هذه القيم هي صفة «ممتاز» يليها «جيد جدا مرتفع» وهكذا حتى نصل إلى أدناها وهي صفة «ضعيف».

والنوع الثالث من المتغيرات هو متغير عددي. والمتغير العددي يتميز بكل ما يتميز به المتغيران الوصفي والترتيبي، أي أنه يصنف، ويُقيم ترتيبا ولكنه بالإضافة إلى ذلك ينبئنا، في مجال الترتيب، بجواب واحد دقيق عن الفارق بين صفة أعلى وصفة أدنى، أو قيمة أعلى وقيمة أدنى. ومع معرفتنا في مثال التقديرات بأن قيمة «ممتاز» أعلى من قيمة «جيد» لو أننا سألنا ما هو الفارق بينهما تماما لما أمكن الإجابة، إذ لا نعلم أي معنى أو جواب محدد لـ «ممتاز ـ جيد». ولكننا في المتغير العددي يمكن الإجابة على مثل هذا السؤال بدقة تامة. لنصنف، مثلا، فصلا من عشرة طلاب، وفقا لمعدلاتهم العامة، ولنفرض أننا وجدنا الجدول التالي:

المعدل العام	91	87	85	73	71	60	45
عدد الطلاب	1	1	1	2	2	2	1

لنرمز للمعدل العام بالرمز Z، فالمتغير Z هو متغير عددي لأن قيمه الممكنة أعداد حقيقية. وقد صنف المتغير Z طلاب الفصل وفق معدلاتهم فظهر معنا سبعة أصناف هي «ذوو المعدل 91» ، ذوو المعدل 87» الخ. وبالطبع يمكن ترتيب الأعداد من الأكبر إلى الأصغر أو بالعكس، بالاضافة إلى أن الفرق بين أي قيمتين محسوب تماما وبدقة. والفارق بين الصنف 91 والصنف 71 هو 20 درجة. والبيانات الإحصائية التي تتضمن قياسات متغير وصفي تسمى بيانات وصفية، وتلك التي تتضمن قياسات متغير ترتيبي تسمى بيانات ترتيبية، أما التي تتضمن قياسات متغير عددي فتسمى بيانات عددية.

ويجب ألا يختلط علينا الأمر عندما نقوم بترميز بيانات وصفية أو ترتيبية وفق نظام رموز عددي معيّن، فلو رمزنا لصفة «سعودي» بالرقم 3 ولصفة «عربي غير سعودي» بالرقم 2 ولصفة «غير عربي» بالرقم1 فإن هذا لا يعني أن بيانا حصلنا عليه يتعلق بجنسيات جماعة من المقيمين في الرياض أصبح بيانا عدديا، ومع أنه سيقتصر على الأرقام 1,2,3 إلا أننا يجب أن نتذكر بأن هذه الأرقام هي رموز لصفات وصفية وأن البيان لا يزال بيانا وصفيا.

وتنقسم البيانات العددية بدورها إلى نوعين، **بيانات عددية منفصلة وبيانات عددية مستمرة**. والبيانات المنفصلة تتضمن قياسات ناتجة عن عملية عد أو تعداد. وعندما نسجل، مثلا، عدد الولادات التي تمت في مستشفى للتوليد في كل يوم من أيام شهر معين سنحصل على بيان إحصائي عددي من النوع المنفصل جميع قياساته أعداد صحيحة. وكذلك الأمر عندما نعد الكريات البيض الظاهرة على منطقة محددة من زجاجة فحص مجهري، وعدد المراجعين الذين زاروا مركزا للرعاية الأولية في يوم معين، وعدد وقوعات الزواج أو الطلاق أو الوفاة خلال فترة محددة في منطقة معينة. وعدد حوادث المرور اليومية في مدينة الرياض الخ. أما النوع الآخر من البيانات العددية وهو البيانات المستمرة فإنها تتضمن قياسات ناتجة عن استخدام جهاز للقياس مثل مسطرة أو ميزان لقياس وزن أو درجة حرارة أو ضغط الدم أو الضغط الجوي، أو اختبار (أورائز) لقياس حاصل الذكاء أو اختبار لقياس معلومات طالب في مقرر معين الخ. ويسمى المتغير العددي الذي تكون قيمه الممكنة من النوع المنفصل أي نحصل عليها بطريقة التعداد، متغيرا عدديا منفصلا، كما يسمى ذاك الذي تكون قيمه الممكنة من النوع المستمر، أي نحصل عليها باستخدام جهاز للقياس، متغيرا عدديامستمرا. ونلاحظ بسهولة أن القيم الممكنة لمتغير عددي منفصل قابلة للعد، بمعنى أنه يمكننا القول إن هذه القيمة هي القيمة الأولى الممكنة تليها القيمة كذا كقيمة ثانية، تليها القيمة كذا كقيمة ثالثة، وهكذا. ونطمئن إلى أننا عند الانتقال من قيمة إلى القيمة التي تليها لم نغفل بينهما أيا من القيم الممكنة للمتغير. فمثلا، لو رمزنا بـ x لعدد الولادات اليومية في مستشفى للتوليد، فإن القيم الممكنة لـ x هي إما صفر، أو 1 أو 2 أو 3 إلخ. ولا يمكن أن يكون هناك نصف ولادة أو ولادة ونصف إلخ. وعندما ننتقل من الصفر

كقيمة ممكنة إلى الواحد كقيمة ممكنة تالية لها، لم نخلف وراءنا أيا من قيم x الممكنة.
وإذا حاولنا تطبيق الفكرة ذاتها في مجموعة القيم الممكنة لمتغير مستمر، أي نحاول عدها، فسنجد أن ذلك مستعص. لنرمز بـ y، مثلا، لطول إنسان ذكر بالغ من رعايا المملكة. فعند قياس طوله بمسطرة مدرجة مرتفعة ومناسبة، كتلك التي نجدها في المستشفيات لقياس الطول، سنجد أن طوله يكافىء نقطة على تدريج المسطرة، لنفرض أن هذه النقطة واقعة بين 171 سم و 172 سم، فطول الرجل، واقع إذا، في مكان ما بين 171 سم و 172 سم. أي أنه يمكن أن يكون أي نقطة من الفترة [171,172] من محور الأعداد الحقيقية. ولو كلف المرء نفسه بعد هذه القيم الممكنة فسيقول إن 171 سم هي قيمة ممكنة أولى ثم يتوقف عاجزا تماما عن تحديد القيمة التي تليها. إذمهما اتخذ عددا قريبا من الـ 171 فبين الـ 171 وبين هذه القيمة، على قربها الشديد من 171، ما لا يحصى ولا يُعد من القيم. ونقول إن مجموعة القيم الممكنة لمتغير مستمر هي مجموعة غير قابلة للعد. وقابلية العد هي الخاصة الرياضية التي نميز بموجبها بين النوعين من البيانات العددية، النوع المنفصل والنوع المستمر.

٩ ـ تدوير الأرقام العشرية ـ أخطاء القياسات

رأينا أن قياس طول شخص يقابل نقطة على المسطرة المدرجة التي نستخدمها في قياس الطول. وهذه النقطة من المسطرة المدرجة (نقطة من محور الأعداد) تقابل أو تمثل عددا حقيقيا هو طول الشخص. ولكن هب أن المسطرة التي نقيس بها مدّرجة بالسنتمتر، وليس فيها تدريج ميليمتر. وكل ما في الأمر أن هناك نقطة تشير إلى منتصف المسافة بين رقم والرقم الذي يليه، شكل (٤)، ولنفرض أن النقطة c على حرف المسطرة هي النقطة المقابلة لقمة رأس الشخص. فالمسطرة بما أوتيت به من دقة في التدريج تخبرنا أن طول الشخص هو عدد حقيقي واقع بين 171 سم و 172 سم إلا أنه أقرب إلى 172 سم منه إلى 171 سم (النقطة C واقعة بعد منتصف المسافة بين 171 و 172) ومن المنطقي جدا، في غياب أية معلومات أخرى، أن نصطلح على القول إن طول الشخص مقرب إلى أقرب سنتمتر هو 172سم. وكذلك النتيجة ستكون لو أن النقطة c وقعت في أي مكان من المنطقة المظللة على المحور، التي تمتد بين 171.5 سم إلى 172.5 سم. ولكن

ماذا لو وقعت c عند الـ 171.5 سم تماما أو عند الـ 172.5 تماما؟ في مثل هذه الحالة نتفق على تقريب الـ 171.5 سم إلى 172 سم، وتقريب الـ 172.5 سم إلى 173 سم. لنفرض الآن وجود تدريج ميلليمتري، فما اصطلحنا عليه سابقا يكافىء ما يلي:

إذا كان القياس 171.5 سم أو 171.6 سم أو 171.7 سم أو 171.8 سم أو 117.9 نعتبره 172 سم وكذلك نعتبره 172 سم إذا كان القياس 172.1 سم أو 172.2 سم أو 172.3 سم أو 172.4 سم. وهذا يُملي علينا، بصورة عامة، القاعدة التالية لتدوير الأرقام العشرية:

167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183

شكل (٤)

قاعدة

لتدوير عدد عشري إلى منزلة معينة ننظر في الرقم الذي يحتل المنزلة التي تليها فإذا كان 5 أو أكثر نضيف واحدا إلى المنزلة المطلوبة ونلغي جميع الأرقام العشرية التي تليها. وإذا كان 4 أو أقل نبقي المنزلة المطلوبة كما وردت ونستغني كذلك عن جميع الأرقام العشرية التي تليها.

مثال (١٥)

كيف تصبح الأعداد التالية بعد تدويرها إلى أقرب جزء من عشرة.

181.253 ،314.0621 ،7.3198 ،0.9808 ،0.0532 ،9.1701

الحـــل

وفقـا للقاعــدة المذكـورة أعلاه، ننظر إلى الـرقم العشري الثاني فـإذا كان 5 أو أكـثر نضيف 1 إلى الرقـم العشري الأول (وهـو الـرقم الذي يحتـل منزلـة الجـزء مـن عشرة) وإذا كان أقل من 5، نحتفظ بالرقـم العشري الأول كما ورد. وهكذا نجد، على الترتيب، 181.3 ،314.1 ،7.3 ،1.0 ،0.1 ،9.2.

مثال (١٦)

فيما يلي عـدد الـزيـارات التي قـام بها المرضى المراجعـون للعيـادات الخارجيـة بالمستشفيـات ومراكـز الرعايـة الصحية الأوليـة في المناطق الأربع عشرة في المملكـة لعام ١٤٠٦هـ: 11168617، 4330131، 4870214، 3028921، 2049754، 4801820، 1577160، 6374554، 6034118 ، 1479722 ، 3875879 ، 3464826 ، 2825761 ، 1793849 .

والمطلـوب التعبير عن هذا البيـان العددي بـآلاف الأشخاص ثم تـدوير الـرقم الناتج إلى أقرب ألف.

الحـــل

التعبير عن البيـان بـآلاف الأشخاص يعني أن وحـدة القيـاس أصبحت «ألف شخص» فكل ألف مراجع يشكلون جماعة واحدة تتضمن ألف شخص. وللتعبير عن هـذه الأعداد بـآلاف الأشخاص يجب أن نقسم كـل عدد على ألف. وتـدوير الأعـداد الناتجة إلى أقـرب ألف يعني تدويرهـا إلى الرقم الذي يحتل منـزلة الآحاد. وهكـذا نجد الأعداد معبرا عنها بآلاف الأشخاص كما يلي:

11168.617، 4330.131، 4870.214، 3028.921، 2049.754، 4801.820، 1577.160، 6374.554، 6034.118 ، 1479.722 ، 3875.879 ، 3464.826 ، 2825.761 ، 1793.849 .

وبعد تدويرها إلى أقرب ألف نجد:

11169، 4330، 4870، 3029، 2050، 4802، 1577، 6375، 6034، 1480، 3876، 3465، 2826، 1794.

ونلاحظ أنه من الأيسر على القارىء متابعة البيان عندما يُعطى بهذا الشكل.

وفي الوقت الذي لا تخضع فيـه قياسات بيان عـددي من النوع المنفصل لأخطاء فإن القياسات في بيان عـددي من النوع المستمر تخضع دائما لخطأ يسمى خطأ القياس. ويعـود السبب في ذلك إلى أننا نستخـدم للوصـول إلى مثل ذلك القياس جهـازا أو أداة

للقياس، ولا يمكن للإنسان أن يبتكر جهازا للقياس لا يخطىء. لقد توصل الإنسان إلى ابتكار أجهزة قياس في علوم الفيزياء والكيمياء وغيرها تقيس بدقة هائلة إلا أنها مع ذلك تخطىء. وبالطبع يضاف إلى هذا السبب أو المصدر مصادر أخرى نذكر منها أن الإنسان الذي يقيس مُعرض أيضا لارتكاب خطأ، ومهما أحسن استخدام الجهاز الذي يقيس به فسيرتكب أيضا نوعا من الخطأ.

وعندما نطلّع على بيان إحصائي عددي من النوع المستمر ينبغي أن نفهم من القياس المقدم لنا شيئين، أولهما فكرة عن مقدار الشيء المقيس وثانيهما درجة الدقة التي يتمتع بها القياس. وإذا قيل لنا إن طول شخص هو 168.7 سم فإن هذا الرقم يعطينا فكرة عن قامة الشخص ولكنه يعطينا أيضا أن دقة هذا القياس تصل إلى أقرب جزء من عشرة من السنتمتر، أي إلى أقرب ميلليمتر. وكقاعدة عامة، يكون الرقم الأخير المعطى على اليمين رقما مشكوكا في صحته. وعندما نقيس، بطريقة علمية، طول شخص ونفيد بأن طوله 168.7 سم فهذا يعني أن الطول الحقيقي لهذا الشخص واقع في مكان ما بين 168.65 سم و 168.75 سم. ولتوضيح الفكرة نقول إننا لو استخدمنا جهازا أكثر دقة لقياس الطول لأعطانا الطول صحيحا حتى منزلة الجزء من مائة، أي حتى الرقم العشري الثاني بعد الفاصلة، وفي هذه الحالة سيكون الرقم العشري الأول بعد الفاصلة صحيحا والشك لا يتطرق إلا إلى الرقم الذي يليه، ولو أن دقة الجهاز سمحت بإعطاء ثلاثة أرقام عشرية بعد الفاصلة أي بدقة تصل إلى أقرب جزء من ألف من السنتمتر، فسيكون الرقمان العشريان الأول والثاني بعد الفاصلة صحيحين ويتطرق الشك إلى الرقم العشري الثالث، وهكذا. وبصرف النظر عن مقدار هذه الأرقام (الرقم العشري الثاني والثالث والرابع الخ. بعد الفاصلة) فإن تدوير العدد الذي نحصل عليه إلى أقرب جزء من عشرة لن يعطينا 168.7 سم إلا إذا كان العدد الذي نقوم بتدويره واقعا بين 168.65 سم، و 168.749999 (ويمكن أن يتكرر الرقم 9 إلى ما لا نهاية له) وذلك وفقا لقاعدة تدوير الأرقام العشرية، ونصطلح هنا، توخيا للسهولة، أن القيمة الفعلية للقياس تقع بين 168.65 سم و 168.75 سم.

وبصورة عـامة، للوصـول إلى الحدين الأدنى والأعلى للقيمة الفعليـة لقياس من النوع المستمـر، معطى على شكل عدد صحيـح (أي مقرب إلى أقرب واحـد صحيح)، نطرح منه 0.5 فنحصل على الحد الأدنى ونضيف إليه 0.5 فنحصل على الحد الأعلى. أما إذا كان القيـاس معطى كعدد عشري فنضع صفرا بعد آخـر رقم معطى في القياس (أي آخر رقم على اليمين بعد الفاصلـة العشرية) ثم نضع الرقم 5 تحت هذا الصفر ونطرح فنحصل على الحد الأدنى ثم نجمع للحصول على الحد الأعلى. (محتفظين بـالفاصلة في موقعها تماما عند الجمع أو الطرح) ونوضح الطريقة بالمثال التالي:

مثال (١٧)

ما هو الحد الحقيقي الأدنى والأعلى لكل من القياسات التالية:

12 سم، 1517 سم، 18.4 سم، 125.05 سم، 34.70 سم، 4.3208 سم؟

الحـــل

الحدود المطلوبة هي على الترتيب:

12.0	12.0
0.5 −	0.5 +
11.5	12.5

فالقياس الأول واقع فعلا بين 11.5 سم و12.5 سم:

وبصورة مماثلة نجـد أن القياس الثاني واقع فعلا بين 1516.5 سم و 1517.5 سم.

ومن أجل القياس الثالث نكتب:

18.40	18.40
0.05 −	0.05 +
18.35	18.45

والقياس الثالث واقع فعلا بين 18.35 سم و 18.45 سم.

وبصورة مماثلة نكتب من أجل القياسات الثلاثة الباقية، على الترتيب،

$$\begin{array}{r} 125.050 \\ 0.005 + \\ \hline 125.055 \end{array} \qquad \begin{array}{r} 125.050 \\ 0.005 - \\ \hline 125.045 \end{array}$$

$$\begin{array}{r} 34.700 \\ 0.005 + \\ \hline 34.705 \end{array} \qquad \begin{array}{r} 34.700 \\ 0.005 - \\ \hline 34.695 \end{array}$$

$$\begin{array}{r} 4.32080 \\ 0.00005 + \\ \hline 4.32085 \end{array} \qquad \begin{array}{r} 4.32080 \\ 0.00005 - \\ \hline 4.32075 \end{array}$$

وفي كل حالة إنما نطرح ونضيف، في الواقع، نصف وحدة دقة. حيث وحدة الدقة هي الواحد في منزلة الرقم المشكوك فيه .

١٠ ـ التناسب الطردي

نقول إن المتغيرين x و y متناسبان طرديا إذا بقيت نسبتهما ثابتة. أي $c = \frac{y}{x}$ ، أو $y = cx$ حيث c عدد ثابت يسمى ثابت التناسب .

لنفرض الآن أن المقدارين x و y يتغيران متناسبين طرديا . ولنفرض أن x تغير من x إلى $x + \Delta x$، وفي مقابل ذلك تغير y من y إلى $y + \Delta y$. ما هي العلاقة بين Δx تغير x و Δy تغير y؟ وللإجابة نفترض أن ثابت التناسب c ، فيكون :

$$\frac{y}{x} = c \, , \, \frac{y+\Delta y}{x+\Delta x} = c$$

ومنه

$$\frac{y+\Delta y}{x+\Delta x} = \frac{y}{x}$$

ومن خواص التناسب يمكن أن نكتب:

$$\frac{y+\Delta y}{x+\Delta x}=\frac{-y}{-x}=\frac{(y+\Delta y)-y}{(x+\Delta x)-x}=\frac{\Delta y}{\Delta x}$$

أي أن $\frac{\Delta y}{\Delta x}=c$ أو $\Delta x = c\Delta x$

فالتغيران في x و y يحافظان دائما على علاقة التناسب الطردي ذاتها القائمة بين x و y. ونلجأ إلى هذه الحقيقة البسيطة في كثير من التطبيقات. فإذا علمنا مثلا أنه عندما ازدادت قيمة x بمقدار 5، ازدادت قيمة y بمقدار 3؛ فكم سيزداد y مقابل زيادة في x مقدارها 7؟ ولحساب المطلوب نكتب:

$$\frac{3}{5}=\frac{\Delta y}{7}$$

ومن خواص التناسب نعلم أن

$$5\times\Delta y=3\times 7$$

$$\Delta y=\frac{3\times 7}{5}=4.2$$

ونلخص عادة هذه العمليات في مخطط بسيط كما يلي:

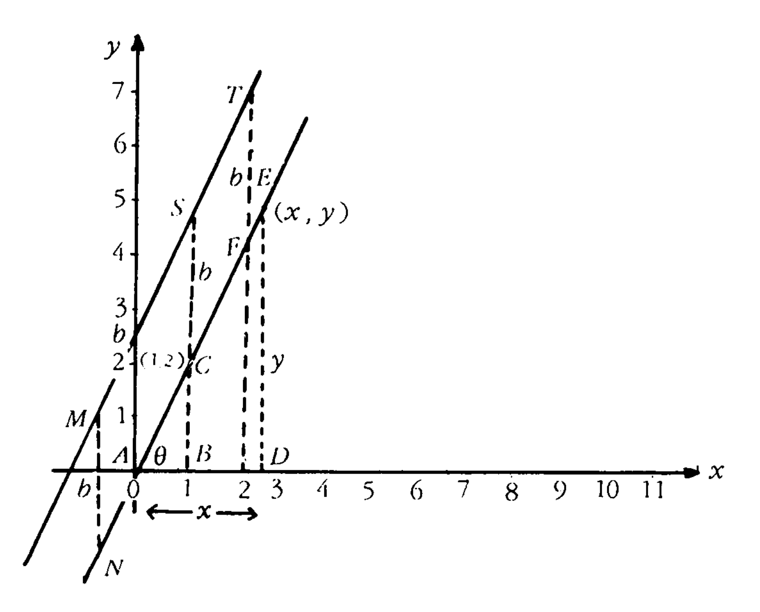
ΔX	ΔY
5	3
7	?

وتكون الزيادة المطلوبة في y مساوية $\frac{3\times 7}{5}=4.2$

١١ ـ معادلة مستقيم

لندرس أولا معادلة مستقيم يمر من مبدأ الاحداثيات. ويتحدد المستقيم تماما عند معرفة نقطتين منه. وفي حالتنا هنا نعلم سلفا أن المستقيم يمر من النقطة (0,0) وهي مبدأ الاحداثيات، فتكفي معرفة نقطة واحدة أخرى لكي يكون ممكنا تحديد معادلة المستقيم. لنفرض أن النقطة (1,2) واقعة على المستقيم فكيف نحدد معادلته؟

إذا رسمنا محورين للاحداثيـات وحددنا النقطة (1,2) ثم وصلنا بينهـا وبين المبدأ نحصل على بيان المستقيم. وعلى هذا البيان نأخذ نقطة ما، نرمز لاحداثياتها بـx وy ، وتكون المعادلة المطلوبة هي علاقة بين x وy . ومن تشابه المثلثين ABC و ADE نكتب:



شكل (٥)

أو
$$\frac{AB}{AD} = \frac{BC}{DE}$$

ومنه
$$\frac{1}{x} = \frac{2}{y}$$
$$y = 2x$$

وهي ذات العلاقة التي تـربط بين مقدارين متنـاسبين طرديـا، حيث ثابت التنـاسب يساوي 2. وتنبغي ملاحظة أن ثابت التناسب 2 يمثل ظل الزاوية θ (حرف يوناني منطوقه ثيتـا) التي يصنعهـا المستقيم مـع الاتجاه الموجب لمحـور السينـات. (انظـر الشكل ٥) ويسمى ظل الزاوية θ ميل المستقيم. وكل مستقيم آخر في المستـوى ميله 2، أي يصنع مع الاتجاه الموجب لمحـور السينات زاويـة تساوي زاويـة المستقيم AE ، سيقطع المحور الصـادي في نقطة غير نقطـة المبدأ. لنفـرض مستقيماً MT موازيـا لـAE ويقطع المحور

الصادي في نقطة $(b, 0)$. فما معادلتـه؟ يمكن الحصول على نقـاط المستقيم MT بإضافة مقدار ثابت b إلى الاحداثي الصادي للنقـاط الموافقة من المستقيم $y = 2x$، وهي النقاط التي تقع على الخط الرأسي نفسه. فإذا أضفنـا إلى الاحداثي الصادي للنقطة N مقدار b حصلنا على M وإذا أضفنا المقدار b إلى الاحداثي الصادي لكل من c و F وجدنا، على الترتيب، S و T. وهذا يعني أن معادلة المستقيم الجديد هي

y (نقطة على المستقيم الجديد) = y (نقطة على المستقيم AE) + b

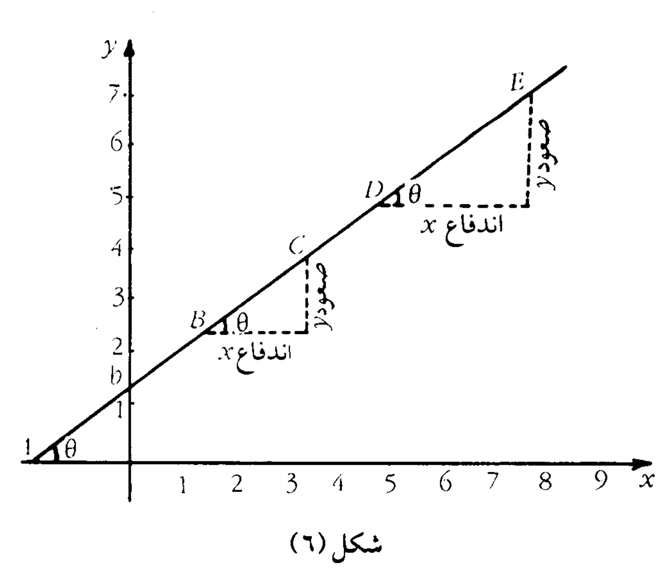
أو

$$y = 2x + b$$

وبصورة عامة نجد أن معادلة المسستقيم AE (انظر الشكل ٦) هي

$y =$ (ظل الزاوية θ) $x + b$

حيث b إحداثي نقطة تقاطع المستقيم مع محور الصادات.



شكل (٦)

ولو انتقلت نقطـة على المستقيم من الوضع B إلى الوضع C فإن x سيتغير بمقدار سميناه «اندفاع x» وسيقابله تغير في y سميناه «صعود y». وكذلك عند انتقال نقطة من

الوضع D إلى الوضع E، فإن x سيتغير بمقدار سميناه «اندفاع x» وسيقابله تغير في y سميناه أيضا «صعود y». ومن خواص الشكل الهندسية نلاحظ بسهولة أن

$$\text{ثابت} = \text{ظل}\ \theta = \frac{\text{صعود}\ y}{\text{اندفاع}\ x}$$

أي أن نسبة تغير y إلى تغير x تبقى ثابتة باستمرار، في حالة مستقيم. وهي الخاصة نفسها التي رأيناها في حالة مقدارين متناسبين طرديا.

وأخيرا، إذا كانت العلاقة بين متغيرين x و y علاقة خط مستقيم قلنا إنها علاقه خطية. وبصورة عامة، معادلة أي مستقيم هي علاقة خطية وبالعكس كل علاقة خطية تمثل مستقيما.

١٢ ـ تصميم الجداول

نحتاج إلى تنظيم نتائج التجارب والمشاهدات العلمية في شكل جداول، وذلك في مختلف ميادين المعرفة. وفي أبسط الحالات نجد جدولا ثنائيا، يتضمن في كل خلية من خلاياه قياسا أو مشاهدة مرتبطة بمتغيرين، ثبتنا كلا منهما عند مستوى معين من مستوياته الممكنة. فنفرض، مثلاً، أن لمتغير x ثلاثة مستويات، سنرمز لها بـ x_1، x_2، x_3؛ وأن لمتغير آخر y أربعة مستويات، سنرمز لها بـ y_1، y_2، y_3، y_4؛ وإذا حصلنا عند كل زوج مختلف من المستويات للمتغيرين x و y على قياس أو مشاهدة، فسيتوفر لدينا اثنا عشر قياسا نضعها في جدول كما في الشكل (٧).

حيث أشير بـ × للقياس وبـــ لمجاميع السطور أو الأعمدة وبـ = للمجموع الكلي. وبصورة عامة، إذا كان عدد مستويات المتغير x مساويا لـ n، وعدد مستويات المتغير y مساويا لـ m فإن عدد خلايا الجدول سيكون $n \times m$.

وفي حال وجود ثلاثة متغيرات x وله n من المستويات، y وله r من المستويات، و z وله q من المستويات، نحتاج إلى تصميم جدول ثلاثي يتضمن $n\ r\ q$ خلية. ونلاحظ بوضوح أننا نحتاج إلى جدول ثنائي مؤلف من $n\ r$ خلية ليستوعب المتغيرين x

x \ y	y_1	y_2	y_3	y_4	المجموع
x_1	×	×	×	×	-
x_2	×	×	×	×	-
x_3	×	×	×	×	-
المجموع	-	-	-	-	=

شكل (٧)

و y . ثم نعيد هذا الجدول q مرة وذلك عند كل مستوى من مستويات المتغير الثالث z . ولتقديم مثال عن تصميم جدول ثلاثي نفترض أن $n = 3$ ، $r = 4$ ، $q = 3$. فنجد جدولا كما في الشكل (٨).

	Z_1			المجموع	Z_2			المجموع	Z_3			المجموع
	x_1	x_2	x_3		x_1	x_2	x_3		x_1	x_2	x_3	
y_1	×	×	×	-	×	×	×	-	×	×	×	-
y_2	×	×	×	-	×	×	×	-	×	×	×	-
y_3	×	×	×	-	×	×	×	-	×	×	×	-
y_4	×	×	×	-	×	×	×	-	×	×	×	-
	-	-	-	=	-	-	-	=	-	-	-	=

شكل (٨)

ومن الواضح أنه يمكن تنظيم الجدول بطريقة ثانية تستوعب فيها الجداول الثنائية المكررة مستويات Z و Y . (انظر الشكل ٩) أو بطريقة ثالثة تستوعب فيها الجداول الثنائية المكررة مستويات X و Z . (انظر الشكل ١٠).

	x_1			المجموع	x_2			المجموع	x_3			المجموع
	z_1	z_2	z_3		z_1	z_2	z_3		z_1	z_2	z_3	
y_1	×	×	×	-	×	×	×	-	×	×	×	-
y_2	×	×	×	-	×	×	×	-	×	×	×	-
y_3	×	×	×	-	×	×	×	-	×	×	×	-
y_4	×	×	×	-	×	×	×	-	×	×	×	-
	-	-	-	=	-	-	-	=	-	-	-	=

شكل (٩)

	y_1			المجموع	y_2			المجموع	y_3			المجموع	y_4			المجموع
	x_1	x_2	x_3		x_1	x_2	x_3		x_1	x_2	x_3		x_1	x_2	x_3	
z_1	×	×	×	-	×	×	×	-	×	×	×	-	×	×	×	-
z_2	×	×	×	-	×	×	×	-	×	×	×	-	×	×	×	-
z_3	×	×	×	-	×	×	×	-	×	×	×	-	×	×	×	-
	-	-	-	=	-	-	-	=	-	-	-	=	-	-	-	=

شكل (١٠)

وعلى سبيل المثال، لنفرض أن لدينا أربعة أنواع من الأبقار نرمز لها $C_1, C_2, C_3,$ C_4 . أخضعناها في محطة للتجارب الزراعية تابعة لكلية الزراعة، إلى ثلاثة أشكال من النظام الغذائي هي N_1, N_2, N_3 . وذلك لدراسة أثر النظام الغذائي في إنتاج الحليب اليومي بالكغ لكل من الأنواع الأربعة. وكانت النتائج كما في الجدول الثنائي التالي :

النظام الغذائي \ أنواع البقر	C_1	C_2	C_3	C_4	المجموع
N_1	25	28	30	35	118
N_2	28	29	31	35	123
N_3	27	28	31	34	120
المجموع	80	85	92	104	361

وإذا فرضنا أن التجربة نفسها قد أجريت في محطة للتجارب الزراعية في أبها وذلك لدراسة أثر عامل البيئة والمناخ. إذا رمزنا لعامل البيئة بـ V فلدينا هنا مستويان V_1 وترمز لبيئة المنطقة الوسطى و V_2 وترمز لبيئة المرتفعات الجنوبية الغربية من المملكة. وأن تجربة أبها أعطت النتائج التالية:

النظام الغذائي \ أنواع البقر	C_1	C_2	C_3	C_4	المجموع
N_1	27	30	29	38	124
N_2	29	33	30	36	128
N_3	30	31	29	38	128
المجموع	86	94	88	112	380

فيمكن جمع هذه النتائج في جدول واحد يلخص العوامل الثلاثة N, C و V. وذلك بأشكال مختلفة، منها، على سبيل المثال، الشكل التالي:

	V_1				المجموع	V_2				المجموع
	C_1	C_2	C_3	C_4		C_1	C_2	C_3	C_4	
N_1	25	28	30	35	118	27	30	29	38	124
N_2	28	29	31	35	123	29	33	30	36	128
N_3	27	28	31	34	120	30	31	29	38	128
المجموع	80	85	92	104	361	86	94	88	112	380

أو يمكن تنظيمه على الشكل التالي :

	V_1				V_2			
	N_1	N_2	N_3	المجموع	N_1	N_2	N_3	المجموع
C_1	25	28	27	80	27	29	30	86
C_2	28	29	28	85	31	33	31	94
C_3	30	31	31	92	29	30	29	88
C_4	35	35	34	104	38	36	38	112
المجموع	118	123	120	361	124	128	128	380

كما يمكن كتابته على الشكل:

	N_1			N_2			N_3		
	V_1	V_2	المجموع	V_1	V_2	المجموع	V_1	V_2	المجموع
C_1	25	27	52	28	29	57	27	30	57
C_2	28	30	58	29	33	62	28	31	59
C_3	30	29	59	31	30	61	31	29	60
C_4	35	38	73	35	36	71	34	38	72
المجموع	118	124	242	123	128	251	120	128	248

تمرين: اقترح أشكالا أخرى.
ويمكن كتابة جدول ثنائي يتضمن العاملين C و V، مثلا، بالجمع فوق مستويات العامل N فنجد:

	V_1	V_2	المجموع
C_1	80	86	166
C_2	85	94	179
C_3	92	88	180
C_4	104	112	216
المجموع	361	380	741

وكذلك يمكن كتابة جدول ثنائي يتضمن العاملين C و N، مثلا، بالجمع فوق مستويات العامل V فنجد:

	N_1	N_2	N_3	المجموع
C_1	52	57	57	166
C_2	58	62	59	179
C_3	59	61	60	180
C_4	73	71	72	216
المجموع	242	251	248	741

وبصورة مماثلة يمكن كتابة جدول ثنائي يتضمن العاملين N و V.

ومن المجاميع الفرعية في الشكل (٧) يمكن تشكيل جدولين ثنائيين. فإذا أخذنا المجاميع الفرعية الواردة في وضع شاقولي نجد جدولا 3×4 يمثل نتائج التجربة لو أنها تضمنت z و y فقط. أي نتائج التجربة لو أننا أغفلنا المتغير x أو جمعنا فوق مستويات X. ويسمى هذا الجدول «الجدول الثنائي $z \times y$»، وكذلك لو أخذنا المجاميع الفرعية الواردة في وضع أفقي نجد جدولا 3×3 يمثل نتائج التجربة لو أنها تضمنت Z و X فقط. أي نتائج التجربة لو أننا أغفلنا المتغير y أو جمعنا فوق مستويات Y. ويسمى هذا الجدول «الجدول الثنائي $z \times x$». ولو أخذنا في الشكل (٧) المجاميع الفرعية الواردة في وضع شاقولي لوجدنا جدولا 3×4 يمثل نتائج التجربة لو أنها تضمنت x و y فقط. أي نتائج التجربة لو أننا أغفلنا المتغير Z، أو جمعنا فوق مستويات Z. ويسمى هذا الجدول «الجدول الثنائي $x \times y$». (ضع جداول ثنائية في مثال الأبقار تمثل $C \times N$، $V \times N$، $V \times C$).

ولتوضيح حالة أربعة متغيرات نأخذ المثال التالي. فلنفرض أن لدينا أربعة متغيرات هي x ويقع في ثلاثة مستويات؛ y ويقع في 3 مستويات؛ Z ويقع في 3 مستويات؛ T ويقع في مستويين. فعندئذ يمكن تصميم جدول $y \times x \times Z \times T$ أبعاده $4 \times 3 \times 3 \times 2$ ويتضمن 72 خلية. هي في الواقع ستة جداول كل منها 3×4. (انظر الشكل ١١).

		Z_1			Z_2			Z_3		
		x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
T_1	y_1	×	×	×	×	×	×	×	×	×
	y_2	×	×	×	×	×	×	×	×	×
	y_3	×	×	×	×	×	×	×	×	×
	y_4	×	×	×	×	×	×	×	×	×
T_2	y_1	×	×	×	×	×	×	×	×	×
	y_2	×	×	×	×	×	×	×	×	×
	y_3	×	×	×	×	×	×	×	×	×
	y_4	×	×	×	×	×	×	×	×	×

شكل (١١)

تمارين الملحق الأول

١) قم بالعمليات الحسابية والجبرية التالية إن أمكن:

ا ـ في الفصل أربعون طالبا وخمسون مقعدا وثلاثة نوافذ. كم طالبا ومقعدا ونافذة في الفصل؟

ب ـ في الغرفة ا أربعون طالبا وخمسون مقعـدا، وفي الغرفة المجاورة ب ثلاثون طالبا وخمسة وأربعون مقعدا. ما هو عدد الطلاب وعدد المقاعد في الغرفتين معا؟

جـ ـ $3xy^2z + 0.5\,xy^2z + 1.2xy^2z - 2.2xy^2z = ?$

د ـ $5x^2yz^2 + 3xyz^2 - x^2yz = ?$

هـ ـ $5\sqrt{3} - 7\sqrt{3} + \frac{3}{2}\sqrt{3} = ?$

و ـ $\sqrt{15} - 3\sqrt{5} = ?$

ز ـ $8\,F(3) + 3\,F(3) - 0.5\,F(3) - F(3) = ?$

ح ـ $7\,F(2) - 3\,F(1) = ?$

ط ـ $\frac{2}{3}+\frac{3}{7}-\frac{5}{21}=?$

ك ـ $\frac{3}{32}+\frac{5}{16}+\frac{1}{2}+\frac{7}{8}-\frac{13}{32}-1=?$

ل ـ $0.5403+1.0279+12.03-3.0101-14.123=?$

٢) في صندوق ثلاث كرات حمر وخمس كرات سود وثماني كرات بيض ما هي النسبة المئوية في الصندوق لكل من الكرات الحمر والكرات السود والكرات البيض؟

٣) في الفصل 25 طالبا من طلاب كلية العلوم و 10 من طلاب الحاسب الآلي وإثنان من الهندسة وطالب من العلوم الصحية. ما هي النسبة المئوية لوجود طلاب كليات العلوم والحاسب الآلي والهندسة والعلوم الصحية في الفصل علما أن مجموع طلاب الفصل 38 طالبا؟

٤) احسب ما يلي: $12.025\times 0.19=?$

$$\frac{4}{5}\times 0.61=? \quad ; \quad \frac{2}{9}\times\frac{7}{5}=? \quad ; \quad \frac{7}{3}\times\frac{7}{9}=?$$

$$130.576\div 1.2 \quad ; \quad 0.7895\div 0.05$$

٥) مستخدما خواص التناسب فيما يلي:

ا ـ أحسب x إذا كان $\frac{x}{5}=\frac{3}{2}$

ب ـ أحسب x و y إذا كان $\frac{x}{2}=\frac{y}{3}$ ، و $x+y=10$.

جـ ـ في صندوق كرات بيض وسود نسبة 2 إلى 1 ، على الترتيب، أحسب عدد الكرات من كل نوع إذا علمت أن الصندوق يتضمن 12 كرة.

د ـ إذا كان $\frac{1}{3}=\frac{x}{y}$ و $32=x^2-y^2$ فاحسب x و y .

هـ ـ إذا كان $\frac{z}{3}=\frac{y}{5}=\frac{x}{2}$ و $x+y+z=30$

فاحسب x ، y ، z .

٦) $A = \{ x :$ طالب في جامعة الملك سعود $x\}$ ؛ $B = \{ y :$ طالب متزوج $y\}$
$C = \{ Z :$ طالب لا يدخن $Z\}$
عبر كلاميا عن $A \cap B$، $B \cap C$، $A \cap C$، $A \cap B \cap C$.

٧) أوجد $A \cup B$ في كل من الحالات التالية:

ا ـ $A = \{a, b, c\}$، $B = \{b, c, \text{d}\}$

ب ـ $A = \{a, b, c\}$، $B = \{a, b\}$

جـ ـ $A = \{\bigcirc, \bigstar\}$، $B = \{+, -\}$

٨) لتكن المجموعة الشاملة S هي مجموعة سكان شبه الجزيرة العربية:
$A = \{a :$ مواطن سعودي $a\}$ ، $B = \{b :$ شخص متعلم $b\}$
$C = \{c :$ شخص مغترب $c\}$
عبر كلاميا عن المجموعات التالية:
$\bar{A}$, $A \cap B$, $\bar{B} \cap \bar{C}$, $B \cup C$, $A \cap B \cap C$, $\bar{A} \cup (B \cap C)$
$B \cup \bar{B}$, $A \cup B$, $B - C$, $B - \bar{C}$.

٩) يمثل الشكل (١٢) المقابل ثلاث مجموعات X ، Y ، Z . ظلل المنطقة التي تمثل المجموعات التالية، كل واحدة في رسم مستقل.

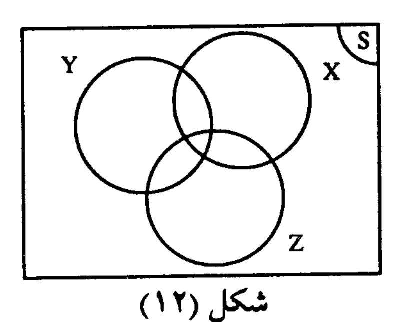
أ ـ $X \cup (Z \cup Y)$

ب ـ $(X \cup Z) \cup Y$ ، وقارن النتائج مع أ.

جـ ـ $(X \cap Y) \cap Z$

د ـ $X \cap (Y \cap Z)$ وقارن النتائج مع جـ .

هـ ـ $X \cap (Y \cup Z)$.



شكل (١٢)

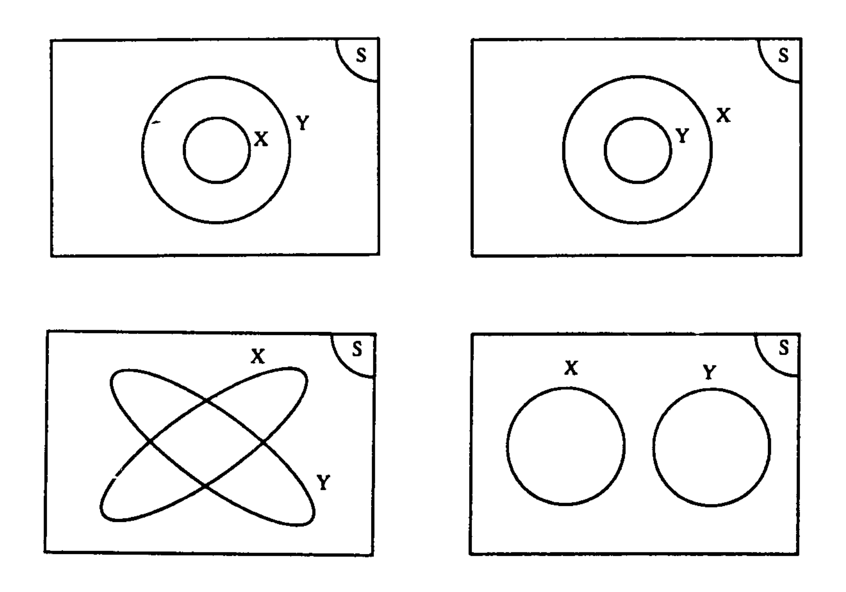
و ـ $(X \cap Y) \cup (X \cap Z)$ وقارن الناتج مع هـ .

ز ـ $X \cup (Y \cap Z)$.

ح ـ $(X \cup Y) \cap (X \cup Z)$ وقارن النتائج مع ز .

لخص النتائج التي حصلت عليها من هذا التمرين بالنسبة إلى قابلية توزيع عملية التقاطع على عملية الاتحاد وتوزيع عملية الاتحاد على عملية التقاطع .

١٠) ظلل $X - Y$ في كل من الأشكال التالية :



شكل (١٣)

١١) في الشكل (١٤) ، المقابل، أكتب المجموعات :

أ ـ X ،

ب ـ Y ،

جـ ـ $X \cap Y$ ،

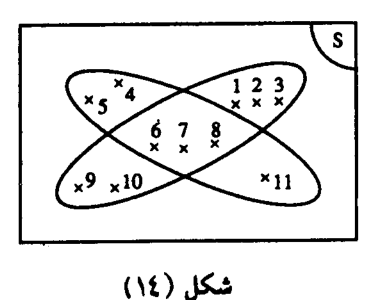
د ـ $X - Y$ ، $X - (X \cap Y)$ ،

هـ ـ $Y - X$ ، $Y - (X \cap Y)$ ،

و ـ $X \cup Y$ ،

ز ـ $(X \cup Y) - X$ ،

لاحـظ أن الناتـج لا يسـاوي Y ، متى يكون الناتج مساويا لـ Y ؟



شكل (١٤)

١٢) اكتب الجـداء الديكـارتي $X \times Y$ إذا كـان $X = \{b, c, d\}$ ؛ $Y = \{m, n, t\}$ ، أكتب أيضا $X \times X$ و $Y \times Y$.

١٣) لتكن الدالة المعرفة بالقاعدة

$$y = \frac{2x+5}{x-3}$$

أ ـ حدد مجموعة تعريف الدالة ومداها ،

ب ـ احسب $f(-1)$ ، $f(0)$ ، $f(4)$ ، $f(5)$ ، $f^{-1}(1)$ ، $f^{-1}(7)$.

١٤) التطبيق $f : Z \to Z$ حيث Z مجموعة الأعداد الصحيحة . معرف كما يلي :

$$f(x) = \begin{cases} 0 & \text{إذا كان } x \text{ عددا يقبل القسمة على 2} \\ 1 & \text{إذا كان } x \text{ لا يقبل القسمة على 2 ويقبل القسمة على 3} \\ 2 & \text{فيما عدا ذلك} \end{cases}$$

ما هي صور الأعداد 8- ، 6- ، 5- ، 0 ، 3 ، 7 ، 11 ، 16 ؟

١٦) لتكن الدالة العددية f المعرفة بالقاعدة :

$$f(x) = \begin{cases} x \ , & x \leq 0 \quad \text{إذا كان} \\ 0 \ , & 0 < x \leq 1 \quad \text{إذا كان} \\ 1 \ , & x > 1 \quad \text{إذا كان} \end{cases}$$

أ ـ أحسب $f(-5)$ ، $f(1/2)$ ، $f(1)$ ، $f(3)$ ، $f(1000)$.

ب ـ أرسم بيان هذه الدالة وعين مداها .

١٧) لتكن الدالة العددية f المعرفة بالقاعدة :

$$f(x) = \begin{cases} x+1 \ , & x < -1 \quad \text{إذا كان} \\ 1 \ , & 1 < x < 2 \quad \text{إذا كان} \\ -2x+3 \ , & x > 3 \quad \text{إذا كان} \end{cases}$$

ا ـ عين مجموعة تعريف هذه الدالة (ساحة الدالة) .

ب ـ أحسب $f(-3)$ ، $f(-2)$ ، $f(3/2)$ ، $f(5)$.

جـ ـ ما هي الصورة العكسية للعدد (2-) .

د ـ أرسم بيان هذه الدالة .

١٨) أكتب بالتفصيل ما تمثله المجاميع التالية :

$$\sum_{i=2}^{6} x_i \quad , \quad \sum_{i=1}^{3} (x_i - 2)^2 \quad , \quad \sum_{i=1}^{4} (x_i + 3)\, x_i \quad , \quad \sum_{i=1}^{4} f_i x_i^2$$

١٩) أكتب كلا من العبارات التالية مستخدما إشارة المجموع $\sum$:

$y_9^2 + y_{10}^2 + y_{11}^2 + y_{12}^2 + y_{13}^2$ ، $x_2^2 + x_3^2 + x_4^2$ ، $x_1 + x_2 + x_3 + x_4$

$r_0 + r_1 + r_2 + r_3 + r_4$ ، $(x_1 - m)^2 + (x_2 - m)^2 + (x_3 - m)^2$ ،

$kn_1 + kn_2 + kn_3 + kn_4 + kn_5$ ، $1^2 + 2^2 + 3^2 + 4^2 + 5^2$ ،

$y_1 + y_2 + y_3 - a - a^2 - a^3$ ، $ay_1 + a^2 y_2 + a^3 y_3 + a^4 y_4$ ،

$x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5$ ، $x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4$

٢٠) إذا كان $x_1 = 3$ ، $x_2 = 2$ ، $x_3 = 1$ ، $x_4 = 0$ ، $x_5 = -3$ حيث

$\sum_{i=1}^{5} x_i = 3$ و $\sum_{i=1}^{5} x_i^2 = 23$. أحسب قيمة كل من العبارات التالية :

(i) باستخدام تعريف $\sum$ ،

(ii) تبسيط العبارة أولا مستخدما خواص $\sum$ ثم حساب القيمة .

أ- $\sum_{i=1}^{5} (x_i + 10)$

ب- $\sum_{i=1}^{5} (2x_i + 3)$

جـ- $\sum_{i=1}^{4} (x_i + x_{i+1})$

د- $\sum_{i=1}^{4} (x_i - x_{i+1})$

هـ- $\sum_{i=1}^{5} (x_i - 1)(x_i + 1)$

٢١) (i) بين أن

$$\sum_{r=0}^{n}\left[(r+1)^2 - r^2\right] = (n+1)^2$$

(أكتب أول حدين وآخر حدين ولاحظ اختصار الحدود السالبة مع الحدود الموجبة).

(ii) بين أن

$$\sum_{r=0}^{n}\left[(r+1)^2 - r^2\right] = 2\sum_{r=0}^{n} r + n$$

(باستخدام خواص Σ).

(iii) من (i) و (ii) بين أن $\sum_{r=0}^{n} r = \frac{1}{2}\, n(n+1)$

٢٢) إذا كانت النقطة $(x, 11)$ واقعة على المستقيم $y = 3x + 5$ فاحسب قيمة x.

٢٣) بين أن النقاط الثلاثة $(2,5)$ ، $(4, 9)$ ، $(1, 3)$ واقعة على استقامة واحدة.

٢٤) في مسح لالتهاب الكبد الفيروسي في مدينة معينة، جرى تسجيل الحالات التي أخبر عنها من المستشفيات، ومن العيادات الطبية، ومن السلطات الصحية المحلية. ويبين الجدول التالي أعداد المرضى الموجودين في مستشفى وغير الموجودين في مستشفى، مصنفين وفقا للجنس، العمر، ولما إذا كانت الحالة من النوع HBSAG أم لا.

أكتب جداول مختصرة تبين تغير نسبة المرضى في المستشفيات مع كل من العمر، الجنس، وحالة الـ HBSAG .

HBSAG سلبي				HBSAG إيجابي				العمر بالسنوات
في مستشفى		ليس في مستشفى		في مستشفى		ليس في مستشفى		
أنثى	ذكر	أنثى	ذكر	أنثى	ذكر	أنثى	ذكر	
0	0	0	0	9	25	42	43	0 - 14
7	16	10	18	20	39	39	41	15 - 29
4	18	3	17	10	21	25	48	≥ 30

٢٥) يتألف فصل الإحصـــاء من 40 طالبـا. صنفوا وفق ثـلاثة متغيرات هـي الجنسية (سعودي، غير سعودي)، والسكن (يعيش في سكن الطلاب، لا يعيش في سكن الطلاب)، والكلية التي ينتسب إليها (علـوم، حاسب آلي، هندسة). إذا علمت أن:

15 طالب سعودي يسكنون في سكن الطلاب ومن العلوم؛ 5 سعوديون لا يسكنون ومن العلوم، 3 طلاب سعوديـون يسكنون في سكن الطلاب ومن الحاسب؛ 2 غير سعوديين يسكنون ومن العلوم، 4 طلاب سعوديون يسكنون في سكن الطلاب ومن الهندسة؛ 1 غير سعــودي يسكن ومن الحاسب، 1 غير سعــودي لا يسكن ومن الحاسب. فاعـرض هذه المعلومات في جـدول علما أن ربع طلاب الفصل من غير السعوديين وأن طلاب الهندسة هم حصرا من السعوديين وجميعهم يعيشون في سكن الطلاب.

الملحق الثاني

بعض الجداول الإحصائية

١ ـ جدول التوزيع الطبيعي المتجمع

$$\phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7743	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9023	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9734	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

جدول توزيع ستيودنت t المتجمع

$$F(t) = \int_{-\infty}^{t} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma(n/2)\sqrt{\pi n}\left(1+\frac{x^2}{n}\right)^{(n+1)/2}}\,dx$$

n \ F	.75	.90	.95	.975	.99	.995	.9995
1	1.000	3.078	6.314	12.706	31.821	63.657	636.619
2	.816	1.886	2.920	4.303	6.965	9.925	31.598
3	.765	1.638	2.353	3.182	4.541	5.841	12.941
4	.741	1.533	2.132	2.776	3.747	4.604	8.610
5	.727	1.476	2.015	2.571	3.365	4.032	6.859
6	.718	1.440	1.943	2.447	3.143	3.707	5.959
7	.711	1.415	1.895	2.365	2.998	3.499	5.405
8	.706	1.397	1.860	2.306	2.896	3.355	5.041
9	.703	1.383	1.833	2.262	2.821	3.250	4.781
10	.700	1.372	1.812	2.228	2.764	3.169	4.587
11	.697	1.363	1.796	2.201	2.718	3.106	4.437
12	.695	1.356	1.782	2.179	2.681	3.55	4.318
13	.694	1.350	1.771	2.160	2.650	3.012	4.221
14	.692	1.345	1.761	2.145	2.624	2.977	4.140
15	.691	1.341	1.753	2.131	2.602	2.947	4.073
16	.690	1.337	1.746	2.120	2.583	2.921	4.015
17	.689	1.333	1.740	2.110	2.567	2.898	3.965
18	.688	1.330	1.734	2.101	2.552	2.878	3.922
19	.688	1.328	1.729	2.093	2.539	2.861	3.888
20	.687	1.325	1.725	2.086	2.528	2.845	3.850
21	.686	1.323	1.721	2.080	2.518	2.831	3.819
22	.686	1.321	1.717	2.074	2.508	2.819	3.792
23	.685	1.319	1.714	2.069	2.500	2.807	3.767
24	.685	1.318	1.711	2.064	2.492	2.797	3.745
25	.684	1.316	1.708	2.060	2.485	2.787	3.725
26	.684	1.315	1.706	2.056	2.479	2.779	3.707
27	.683	1.314	1.703	2.052	2.473	2.771	3.690
28	.683	1.313	1.701	2.048	2.467	2.763	3.674
29	.683	1.311	1.699	2.045	2.462	2.756	3.659
30	.683	1.310	1.697	2.042	2.457	2.750	3.646
40	.681	1.303	1.684	2.021	2.423	2.704	3.551
60	.679	1.296	1.671	2.000	2.390	2.660	3.460
120	.667	1.289	1.658	1.980	2.358	2.617	3.373
∞	.674	1.282	1.645	1.960	2.326	2.576	3.291

ثبت المصطلحات

- عربي - إنجليزي
- إنجليزي - عربي

أولا : عربي - إنجليزي

ا

اتحاد	Union
احتمال شرطي	Conditional probability
إحصاء	Statistics
وصفي	Descriptive statistics
اختبار فرضية	Hypothesis testing
ارتباط	Correlation
استقلال	Independence
أصلي	Original
انحراف	Deviation
متوسط	Mean deviation
معياري	Standard deviation

ب

بيانات	Data
ترتيبية	Ordinal data
عددية	Interval value data
وصفية	Nominal data
بياني	Graphic

ت

تباين	Variance
تجربة	Experiment
ترتيب تصاعدي	Ascending order
طبيعي	Natural order
ترميز	Coding
تفرطح	Kurtosis
تقاطع	Intersection
تقدير	Estimate
بفترة	Interval estimation
تقريب	Approximation
الثنائي بالطبيعي	Normal approximation to binomial

تكرار	Frequency
تكرارات مستقلة	Independent trials
تكرار متجمع	Cumulative frequency
تمثيل بياني	Graphic presentation
توزيع	Distribution
احتمالي منفصل	Discrete probability distribution
التكرار	Frequency distribution
ثنائي	Binomial distribution
ستيودنت (التوزيع t)	Student's t-distribution
طبيعي	Normal distribution
طبيعي معياري	Standard normal distribution
فوق الهندسي	Hypergeometric distribution
المعاينة	Sampling distribution
توقع	Expectation
رياضي	Mathematical expectation

ج

جدول التكرار	Frequency distribution

ح

حادثة	Event
حادثتان متنافيتان	Mutually exclusive events
منفصلتان	Disjoint events
حادثة بسيطة	Simple event
خالية	Empty event
مركبة	Compound event
حجم العيّنة	Sample size
حد الشقة الأدنى	Lower confidence bound
حوادث مستقلة	Independent events

ر

ربع أول	First quardrant
ربيع أدنى	Lower quarter
أعلى	Upper quarter

ع

عشوائي	Random
عُشير	Decile
عملية	Operation
عنصر	Element
عيّنة	Sample

ف

فترة ثقة	Confidence interval
فرضية	Hypothesis
إبتدائية	Null hypothesis
فضاء	Space
العيّنة	Sample space

ق

قانون بايز	Bayes law
قياس	Measurement
قيمة مطلقة	Absolute value

ك

كشف (تفتيش)	Inspection
بطريقة العينة	Sampling inspection



مبدأ العدّ	Counting Principal
متباينة	Inequality
متغيّر	Variable
عشوائي متصل (مستمر)	Continuous random variable
عشوائي منفصل	Discrete random variable
متممة (مكملة)	Complement
متوافقات (توافيق)	Combinations
متوسط	Mean
الانحراف	Mean deviation
حسابي	Arithmatic mean
عيّنة	Sample mean
مرجّح (موزون)	Weighted mean
مجموعة	Set
جزئية	Subset
شاملة	Universal set
مخطط الانتشار	Scatter diagram
ڤن	Venn - diagram
مدى	Range
ربيعي	Interquartile range
مدرج	Hystogram
التكرار	Frequency hystogram
مساواة	Equality
مستوى الدلالة	Level of significance
مسلمات الاحتمال	Axioms of probability
مشاهدة (ملاحظة أو قياس)	Observation
مضلع التكرار	Frequency polygon
المتجمع	Cumulative frequency polygon
معامل	Coefficient
ارتباط	Correlation coefficient
تغير	Coefficient of variation
ثقة	Confidence coefficient
معاينة	Sampling
طبقية	Stratified sampling
معياري	Standard
مقاييس التشتت	Measures of dispersion
النزعة المركزية	Measures of central tendency
مقدمة	Introduction
مقياس	Measure
منحنى	Curve
التكرار	Frequency curve
طبيعي	Normal curve
منطقة حرجة	Critical region
منفصل	Discrete
منوال	Mode



نتيجة	Outcome
نظرية النهاية المركزية	Central limit theorem
نموذج	Model
الاحتمالات المتساوية	Equal probability model



وسيط	Median
وصف بيانات	Description of data



ثانيًا : إنجليزي - عربي

A

Absolute value	قيمة مطلقة
Approximation	تقريب
Arithmatic mean	متوسط حسابي
Ascending order	ترتيب تصاعدي
Axioms of probability	مسلمات الاحتمال



Bayes law	قانون بايز
Binomial distribution	توزيع ثنائي



Central limit theorem	نظرية النهاية المركزية
Coding	ترميز
Coefficient	معامل
of variation	معامل التغيّر
Combinations	متوافقات (توافيق)
Complement	متممة
Compound event	حادثة مركبة
Conditional probability	احتمال شرطي
Confidence coefficient	معامل الثقة
interval	فترة ثقة
Continuous random variable	متغيّر عشوائي متصل مستمر
Correlation	ارتباط
coefficient	معامل ارتباط
Counting principle	مبدأ العدّ
Critical region	منطقة حرجة
Cumulative frequency	تكرار متجمع
polygon	مضلع التكرار المتجمع
Curve	منحنى



Data	بيانات
Deciles	عشيرات
Description of data	وصف بيانات
Descriptive statistics	إحصاء وصفي
Deviation	انحراف
mean	متوسط الانحراف

English	العربية
Discrete	منفصل
probability distribution	توزيع احتمالي منفصل
random variable	متغيّر عشوائي منفصل
Disjoint events	حادثتان منفصلتان
Distribution	توزيع

E

English	العربية
Element	عنصر
Empty event	الحادثة الخالية
Equality	مساواة
Equal probability model	نموذج الاحتمالات المتساوية
Estimate	تقدير
Event	حادثة
Expectation	توقع
Experiment	تجربة

F

English	العربية
First quardrant	الربع الأول
Frequency	تكرار
curve	منحنى التكرار
distribution	توزيع التكرار
histogram	مدرج التكرار
polygon	مضلع التكرار
table	جدول التكرار

G

English	العربية
Graphic	بياني
presentation	تمثيل بياني

H

English	العربية
Histogram	مدرج
Hypergeometric	فوق الهندسي
distribution	توزيع فوق الهندسي
Hypothesis	فرضية
testing	اختبار فرضية

I

English	العربية
Independence	استقلال
Independent events	حوادث مستقلة
trials	تكرارات مستقلة
Inequality	متباينة
Inspection	كشف (تفتيش)
Interquartile range	المدى الربيعي
Intersection	تقاطع
Interval estimation	تقدير بفترة
valued data	بيانات عددية
Introduction	مقدمة

K

English	العربية
Kurtosis	تفرطح

L

English	العربية
Level of significance	مستوى الدلالة
Lower confidence bound	حد الثقة الأدنى
quartile	الربيع الأدنى

M

English	العربية
Mathematical expectation	توقع رياضي
Mean	متوسط
deviation	الانحراف المتوسط
Measure	قياس (مقياس)
Measurement	قياس

المئينات	Percentiles
متبادلات (تباديل)	Permutations
نقطة	Point
تقدير نقطي	estimation
توزيع بواسون	Poisson distribution
مجتمع	Population
عرض البيانات	Presentation of data
احتمال	Probability
خواص	Properties
نسبة	Proportion

Q

الربيعات	Quartiles

R

عشوائي	Random
تجربة عشوائية	experiment
رقم عشوائي	number
عيّنة عشوائية	sample
متغيّر عشوائي	variable
مدى	Range
ارتباط الرتب	Rank correlation
نسبة	Ratio
البيانات الخام	Raw data
حقيقي	Real
أعداد حقيقية	numbers
محور الأعداد الحقيقية	numbers axis
تكرار نسبي	Relative frequency
قاعدة	Rule
قواعد الاحتمال	Rules of probability

مقاييس	Measures
مقاييس النزعة المركزية	of central tendency
مقاييس التشتت	of dispersion
وسيط	Median
منوال	Mode
نموذج	Model
حادثتان متنافيتان	Mutually exclusive events

N

ترتيب طبيعي	Natural order
بيانات وصفية	Nominal data
تقريب التوزيع الثنائي بالتوزيع الطبيعي	Normal approximation to binomial
المنحنى الطبيعي	curve
توزيع طبيعي	distribution
الفرضية الابتدائية	Null hypothesis

O

مشاهدة (ملاحظة) (قياس)	Observation
عملية	Operation
بيانات ترتيبية	Ordinal data
أصلي	Original
نتيجة	Outcome

P

وسيط (معلمة)	Parameter
جزئي	Partial
تجزئة	Partition
نسبة مئوية	Percentage

معاينة طبقية	Stratified sampling
توزيع ستيودنت (توزيع - t)	Student's t-distribution
مجموعة جزئية	Subset

U

اتحاد	Union
مجموعة شاملة	Universal set

V

متغيّر	Variable
تباين	Variance
مخطط فن	Venn-diagram

W

المتوسط المرجّح	Weighted mean
أوزان	Weights

S

عيّنة	Sample
متوسط عيّنة	mean
حجم العيّنة	size
فضاء العيّنة	space
معاينة	Sampling
توزيع المعاينة	distribution
الكشف بطريقة العيّنة	inspection
مخطط الانتشار	Scatter diagram
مجموعة	Set
حادثة بسيطة	Simple event
فضاء	Space
معياري	Standard
انحراف معياري	deviation
التوزيع الطبيعي المعياري	normal distribution

المراجع

Cramer, H. *Mathematical Methods of Statistics.* Princeton: Princeton University Press, 1961.

Clarke, G.M. and Cooke, D. *A Basic Course In Statistics.* England, Bath: Edward Arnold, 1982.

Campbell, R.C. *Statistics for Biologists, 2nd* ed. England: Cambridge Univ. Press, 1974

Dunn. O.J. *Basic Statistics: A Primer for the Biomedical Science, 2nd* ed. New York: John Wiley, 1977.

Daniel, W.W. *Biostatistics- A Foundation for Analysis in the Health Sciences.* Singapore: John Wiley, 1987

Feller, F. *An Introduction to Probability Theory and Its Applications,* Vol. I, 3*rd* ed. New York: John Wiley, 1967.

Freund, J.E. *Modern Elementary Statistics, 5th* ed. New Jersey: Prentice Hall, 1979.

Hodge, S.F. and Seed, M.L. *Statistics and Probability, 2nd* ed. Edinburgh: Blackie & Chambers, 1977.

Huntsburger, H. *Elements of Statistical Inference.* Boston: Allyn and Bacon Inc., 1981

Handel, D.J. *Inroductory Statistics for Sociology.* New Jersey Prentice-Hall Inc., 1978.

Kendall, M. and Stuart, A. *Advanced Theory of Statistics,* vol. 1, 4*th* ed. London: Charles Griffin & Company, 1977.

Levin, J. *Elementary Statistics in Social Research.* New York: Harper and Row Publishers, 1977.

Mendenhall, W. *Introduction to Probability and Statistics, 6th* ed. Boston: Duxbury press, 1983

Osborn, J. F. *Statistical Exercises in Medical Research.* Oxford: Blackwell Scientific Publications, 1979.

أنيس كنجو. الإحصاء وطرق تطبيقه في ميادين البحث العلمي (الجزء الأول). بيروت: مؤسسة الرسالة، ١٩٧٩م.

كشاف الموضوعات

ح

خ

د

ر


العبيكان
Obeikan
(01) 4983392